



Toward an Ideological Common Space: Extending Bonica's CFscores to the Citizen Level

Seth Babson Warner^{1,2} 

Accepted: 30 May 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract

Bonica's (Am J Polit Sci 58(2):367–386, 2014) campaign finance-based ideology scores, or CFscores, create an ideological common space that allows researchers to compare a wide variety of actors. Because relatively few citizens donate to candidates, however, the public is not well represented in this common space. This paper addresses that gap. It uses random forest machine learning on data from the 2012 Cooperative Congressional Election Study to impute CFscores for respondents who did not donate to candidates, based on how their policy views compared to those who did. These new scores are robust to differences in issue importance between donors and non-donors, and they outperform other ideological measures in predicting vote choice. The scores are then applied to a substantive exercise. Past research shows that extreme candidates for governor are penalized more by voters than those in lower-profile races. The implied mechanism—that vote choice for governor is more ideologically-driven—can be directly tested with imputed CFscores, since they uniquely allow comparisons between voters and candidates across races. An analysis of voting behavior in 2012 confounds expectations. Ideology appears to factor no more into vote choice for governor than for US House. These novel findings underscore the value of extending CFscores to non-donating survey respondents, and while current efforts are limited by data availability, this study offers encouragement and a roadmap to that end.

Keywords Ideology · CFscores · Measurement · Vote choice · Political methodology · Campaign finance · State politics · Governors

Ideology is central to the study of political behavior, and its measurement is a key methodological challenge. Scholars have devoted considerable time and energy

✉ Seth Babson Warner
sbw59@psu.edu

¹ The Pennsylvania State University, State College, PA, USA

² University of North Florida, Jacksonville, FL, USA

toward measuring the ideologies of actors as disparate as legislators (McCarty, Poole, and Rosenthal 2005), judges (Bailey, 2016), bureaucrats (Bonica et al., 2015a, 2015b), and the public at large (Berry et al., 2010; Enns & Koch, 2013).

More difficult still has been the challenge of comparing ideology *between* groups, because each set of actors produces different observable behaviors. Citizens respond to survey items, judges rule on cases, legislators cast floor votes, and executives sign or veto bills. The distinctions between these behaviors—and the contexts they are observed in—make it challenging to place different types of actors on a common ideological scale.

Campaign finance-based ideology scores (CFscores), developed by Adam Bonica (2014), promise a solution to this “common space problem.” Bonica compiled a database of over 103 million donations and used patterns of “who donates to whom” to estimate the ideological preferences of donors and candidates alike. The fact that most actors in the political system either give or receive campaign contributions allows them to be placed on the same ideological scale. CFscores have been used to compare the preferences of legislators and their donors (Burke et al., 2021; Kujala, 2020); judges and executives (Bonica & Woodruff, 2015); corporate boards (Bonica, 2016); think tanks (Lerner, 2017); political operatives (Martin & Peskowitz, 2018); bureaucrats (Bonica et al., 2015a, 2015b); and even medical doctors and law clerks (Bonica et al., 2017, 2020).

However, the large common space created by CFscores does not yet include the key group in the study of democratic government: the public. While some citizens donate to candidates and have CFscores, the vast majority do not, and the literature lacks a way of placing non-donating citizens onto the ideological scale that CFscores create.

In this paper, I present a method of extending CFscores to non-donating citizens, using data from Hill and Huber (2017) and the 2012 Cooperative Congressional Election Study (CCES). Hill and Huber commissioned YouGov to pair donors who responded to the CCES with their CFscores in the Database on Ideology, Money, and Elections (DIME, Bonica, 2014). The merged dataset allows me to compare donors and non-donors based on how they replied to the same policy questions, and use those replies to impute CFscores for citizens who did not contribute to candidates.

In my analysis, I show that issue positions can predict CFscores with a high degree of accuracy. I demonstrate the viability of comparing the policy views of donors and non-donors, because despite their differences, they draw on largely the same issues to form their left–right ideological preferences. Then, I show that imputed CFscores for non-donors look and behave as they should, and outperform other methods of measuring ideology in predicting behavior.

Finally, I apply these scores to a substantive discussion on the role of ideology in gubernatorial elections, where researchers have remarked on high electoral penalties for ideological extremism and the preponderance of moderate winners (Caughey & Warshaw, 2019; Warner, 2022). Imputed CFscores allow me to place incumbents, challengers, and voters alike on a common space across elections at multiple levels of government in the 2012 cycle. This analysis yields a surprising result: ideology does not appear to have any greater impact on gubernatorial vote choice than

on votes in lower-profile elections for US House. This in turn points the literature toward alternative explanations for the relative success of moderate gubernatorial candidates, such as turnout effects or party strength.

Driving this exercise is a sense of excitement that, with CFscores, researchers are on the cusp of a “true” ideological common space, one that directly compares the preferences of most actors in the political system. But we are not there yet, and indeed, the approach I lay out is currently only possible with the 2012 CCES. In the conclusion, I describe what future data collection efforts may look like. While the work involved would be far from trivial, the good news is that random forest machine learning—which drive the models used in this paper—can create measures comparable in quality to those seen here with relatively small amounts of data.

The structure of the paper is as follows. I begin with an overview of past efforts to measure the ideology of citizens and elites on a common scale, and then describe the data used to create imputed CFscores. I discuss the comparability of donor and non-donor issue positions, and show how issue positions can at once predict donor CFscores, and be used to impute valid measures of ideology for non-donors. Then, I present my substantive analysis of gubernatorial elections before concluding with a discussion of the next steps. Ultimately, I argue that expanding the CFscore common space to citizens would mark an impressive methodological achievement, and open new avenues of inquiry for scholars interested in representation, responsiveness, and electoral behavior.¹

Approaches to Creating an Ideological Common Space

Almost since the advent of empirical political science, researchers have sought to compare the preferences of citizens and elites, as such comparisons are required to study many key questions in the discipline. To study the policy influence of citizens relative to other groups or actors, for example, one must measure their preferences in a comparable manner. Likewise, to assess the extent to which voters select ideologically similar candidates, one must be able to place citizens and elites on comparable ideological scales. As Lo et al., (2014, 205) write, “research on elections and party competition is unthinkable without *comparable* measures of the ideological positions of voters” and other actors (emphasis added).

Three approaches to making such comparisons stand out as commonplace. The first uses citizen perceptions of their own and other actors’ ideology as prompted by a survey question. Major surveys such as the ANES Time Series Study and the Cooperative Election Study ask respondents to place themselves, political parties, candidates, and institutions (e.g., the Supreme Court) on a scale from “very liberal” to “very conservative,” with “moderate” as the midpoint. Using this approach allows researchers to study the correlates and consequences of perceived ideology, which

¹ Code and data to replicate this study, as well as imputed CFscores for respondents to the 2012 CCES, have been posted online at <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/GHQKSW>.

more directly affects citizen behavior than “objective” ideology (Granberg & Brown, 1992). Moreover, scholars can infer the objective ideology of elites by drawing on the placements of multiple respondents (Aldrich & McKelvey, 1977). The approach has drawbacks, however. Democratic accountability depends on outcomes reflecting citizens’ true preferences, not just what is perceived, and citizen placements can be systematically biased (King et al., 2004).

A separate approach involves asking citizens to register an opinion on topics that elites are addressing. The “direct comparison” approach allows researchers to assess whether citizens get what they want out of government, and how they respond to elite behavior. To the former, Miller and Stokes (1963) famously aggregated survey responses on social welfare policy, foreign involvement, and civil rights to the US House district level, and found that Members of Congress were only substantially responsive to their citizens on the topic of civil rights. More recently, Gilens and Page (2014) estimated support for 1779 different issues among the wealthy, the mass public, and interest groups, to see whose preferences were most determinative of policy adaptation (they conclude that the rich win, cf. Enns, 2015). In the reverse direction, scholars have used direct comparison to study how aware and responsive voters are to elite behavior. Lenz (2009) highlighted the importance of elite cues in opinion formation, showing that as elites pay more attention to an issue, voter preferences shift to align with those of copartisan leaders. Jessee (2009) and Shor and Rogowski (2018) compare survey responses to the policy positions of candidates for president and US House, respectively, to present evidence of ideological voting in these elections.

The direct comparison approach has become a prominent method of comparing citizen and elite preferences. Its drawback, however, lies in a lack of scalability. A large amount of effort may go into identifying survey responses that reflect citizen views on a set of topics, and outcomes such as roll-call votes that reflect elite behavior on them. Even then, such an effort would only allow for a static comparison between two actors. Modeling dynamic effects over time or incorporating more than two sets of actors (e.g., executives, legislators, donors, and citizens) would constitute a large data collection task using the direct comparison approach.²

A third method recovers this scalability by creating measures that summarize actors’ left–right ideologies into a single “ideal point.” Barberá (2015), for example, uses patterns of who follows whom on Twitter to estimate the ideologies of various users on that platform, from ordinary citizens to elected officials and media outlets. Tausanovitch and Warshaw (2013) use survey responses and Census data on the demographics of different geographies to estimate constituent ideology in states, cities, and state legislative districts. These and similar estimates have then been used to assess political responsiveness at various levels of government (Caughy & Warshaw, 2019; Tausanovitch, 2019; Tausanovitch & Warshaw, 2014).

² There are also other, more minor drawbacks. Error in the survey instruments used to measure citizen preferences may make the comparisons in this approach less than one-to-one. Further, to be generalizable, the comparisons being drawn between citizens and elites should be roughly representative of the political agenda.

Bonica's (2014) CFscores are another example of ideal point estimates, and they stand out for their versatility. They provide measures of ideology for every candidate for elected office, regardless of the outcome or branch of government. These measures are also dynamic, updated for every new election cycle in which a politician runs for office.

Drawing on this versatility, scholars have used CFscores to study a variety of relationships, such as the relative influence of executives over appointments and policy outcomes (Bonica et al., 2015a, 2015b; Cooper et al., 2016a, 2016b; Warner, 2022), the effect of candidate extremism on vote behavior (Utych, 2020; Miller, 2022), and the role of the political donors in exacerbating elite polarization (Kujala, 2022; Kilborn and Vishwanath, 2022).

However, there is one group that CFscores have yet to be applied to, and it is perhaps the most important in the study of American democracy. Citizens only receive a CFscore if they donate to candidates, and the vast majority of Americans do not. As such, the "voice of the people" cannot easily be incorporated into models of political decision-making that use CFscores, unless a methodological bridge is built that allows non-donating citizens to be placed on the same scale.

Extending Bonica's CFscores to Non-donors to Create an Ideological Common Space

As described in the previous section, CFscores allow scholars to make one-to-one comparisons across a variety of distinct actors, from legislators to executives, bureaucrats to judges, and pressure groups to individual donors. But at present, they fall short of providing a *universal* common space because CFscores are not assigned to citizens who do not donate to candidates.

This can be rectified, however, if citizens can be compared to other actors who have CFscores on a dimension that sufficiently explains variation in that measure. Here, I propose to compare non-donating citizens to donors based on their respective policy views on the 2012 CCES. While a naïve comparison shows that these groups differ in their demographic and political characteristics, a more advanced analysis reveals that policy positions map onto self-described ideology in largely the same way (i.e., the same issues are important to both donors and non-donors). This allows me to create and validate an imputed measure of CFscores for non-donors, which I then apply to a substantive discussion of spatial voting in Congressional elections.

Comparing Donors to Non-donors in the 2012 CCES

To impute CFscores for non-donating citizens, the first step is to find a population that I can compare them to. Fellow citizens who do, in fact, donate to candidates stand out as an obvious point of comparison. However, these two groups differ in systemic ways. If this affects the way their policy preferences map onto their left–right ideology, then it may result in bias in the imputed measures.

Table 1 Demographic and political differences between donors and non-donors in the 2012 CCES

	Donors	Non-donors	Student's <i>t</i>	p-value
<i>Demographic</i>				
Mean age	61.7 years	51.4 years	56.7	0.00
Family income > 100 k	37.3%	15.4%	27.6	0.00
Bachelor's degree	63.6%	32.6%	41.4	0.00
<i>Political</i>				
Moderate	22.3%	33.3%	- 16.7	0.00
Strong partisan ID	66.6%	41.6%	33.7	0.00
Follow news "most of time"	90.4%	51.7%	78.3	0.00
Knowledge of House and Senate majority party	89.8%	48.8%	81.0	0.00

Surveys of donors have long been used to evaluate how this group differs from the general public (Brown et al., 1980; Grant & Rudolph, 2002), and recent efforts have highlighted the distinctive nature of donors in a polarized and digitized era. Hill and Huber (2017), whose data I use in this study, commissioned YouGov to identify donors who responded to the 2012 CCES. Relative to the rest of the sample, donors tended to be older, wealthier, and more educated than non-donors, as well as more likely to vote in primary elections and hold ideologically extreme positions. Other studies show that donors prefer candidates who are party-typical ideologically, neither too extreme nor moderate (Gooch & Huber, 2020), and that Democratic donors are more socially liberal—and Republican donors more economically conservative—than their party's rank-and-file (Broockman & Malhotra, 2020).

Given the clear differences in donors and non-donors, how can we be sure that comparing their issue positions is a valid way of imputing CFscores? To explore this question, and conduct the analyses that follow, I turn to the 2012 CCES paired with replication data from Hill and Huber's (2017) study of political donors. The 2012 CCES was an online survey of about 54,535 respondents fielded ahead of that year's midterm elections. The authors commissioned YouGov, which administered the survey, to match respondents who donated to candidates with their CFscores in the Bonica (2014) database. In total, 4432 respondents were identified with donors and paired with CFscores. In order to maintain their anonymity, YouGov added random noise from a uniform distribution of $[-0.1, 0.1]$ to each CFscore.

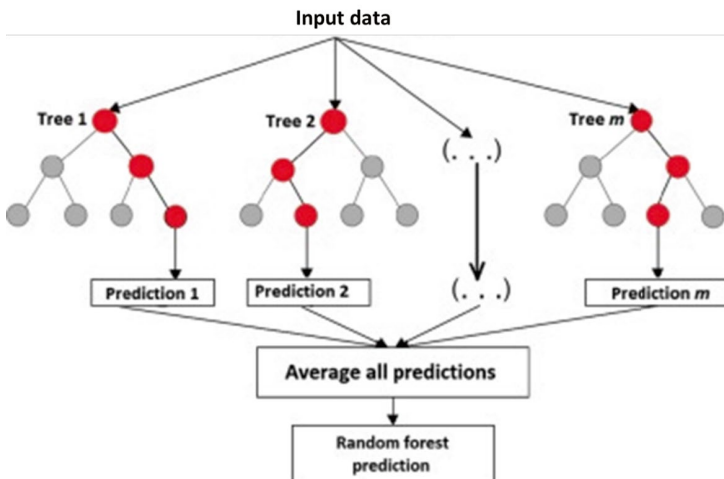
The summary statistics in Table 1 present some examples of how donors and non-donors differ in the survey sample. As Hill and Huber (2017) wrote previously, donors as a group are older, wealthier, and more educated than non-donors, and these differences can sometimes be profound. The percentage of donors with household incomes over \$100,000, for example, is 2.5 times the percentage of non-donors, and nearly twice as many hold bachelor's degrees. These differences continue to politics, where donors are more likely to claim a "strong" partisan affiliation, say that they follow the news "most of the time," and know which parties controlled the House and Senate.

From these comparisons, it is clear that donors and non-donors are systemically different groups. However, do those differences extend to the way their issue positions map onto ideology?

I respond to this question in two ways, one simple and the other more complex. For the former, I run two OLS regressions that use the 25 CCES policy questions to explain donors' and non-donors' self-placements on the survey's five-point ideological scale. Then, I correlate the regression coefficients between models. This yields an association of $r=0.99$. Although simple, this analysis tells me that any given policy view is likely to have a similar effect on a citizen's ideological self-placement, whether or not they donate to political campaigns.

However, to more thoroughly model the process by which issue positions map onto ideology, I turn to random forest machine learning. Random forest modeling is a method that uses covariates to predict some outcome by taking the average prediction of a collection of decision trees (many trees make a forest). At the start of each tree, the algorithm takes a random subset of data and covariates, and "splits" the observations based on the covariate that explains the outcome best. This sorts the observations into two separate nodes, where again, random covariates are drawn, and the observations in each node split on the covariate that explains the most variation in the outcome among them. The process continues in this way until a user-set limit on the number of nodes (or observations per node) is reached. At this point, the mean value of the outcome variable for observations in each node is taken as the tree's prediction for those observations. After all trees have been run, the forest's final predictions are the mean of all trees' estimates for each observation. This process is visualized in Fig. 1.

Random forests have two main benefits in this context. First, they provide a *variable importance vector*, which tells us how much variation in outcome each covariate explains across all levels of all trees. If the same issue positions explain



Note: Used and adapted with permission from Afzal et al. (2020).

Fig. 1 Predicting CFscores from survey questions using a random forest. Note: Used and adapted with permission from Afzal et al. (2020)

ideology for both donors and non-donors, it will give us confidence that they are comparable groups for our purposes. Second, random forests are good at capturing multidimensionality. When researchers use factor or principal component analysis to convert issue positions into a left–right scale, they assume the same data-generating process for every respondent in the data. In reality, however, each person draws on their own, distinct set of issue positions to determine their summary left–right ideology (Cahoon et al., 1976). Random forests can capture this by identifying the most important covariate for observations at each node, which reflects how some issue positions have more impact on ideology for some people than others (Bonica, 2018). It also allows me to incorporate traits that may distinguish donors from non-donors—such as age, income, or news attention—that may condition how issue positions map onto ideology, leaving only the smaller difference of whether or not an individual contributed to a political campaign.

To assess whether issue positions map onto ideology similarly for donors and non-donors, I collect the variable importance vectors from three random forest models trained on data from the 2012 CCES.

In the first model, I use donors' responses to the 25 policy questions to predict their CFscores. In the second model, I use the same issue positions to predict how donors placed themselves on the five-point scale of ideology. Then, in the third model, I again use issue positions to predict how *non-donors* place themselves on this scale.

Variable importance vectors from the first two models allow us to ascertain whether similar issues inform CFscores and ideological self-placements. A high correlation between these vectors serves as validation that CFscores capture latent ideology. Then, variable importance from the third model—which focuses on non-donor ideology—can be compared to that of the others, to determine the extent to which similar issues drive donor and non-donor ideology.

Table 2 presents these correlations between the variable importance vectors for each of these three models. The outcome variables are indicated in the first row and column. In comparing the vectors for the donor CFscore and self-placement models, the correlation of $r=0.95$ indicates that almost the same variables explain these two measures of donor ideology, providing assurance that CFscores viably speak to a donor's left–right orientation.

The other comparisons show that the same issues drive donor and non-donor ideology. This is most notable when the same measure is used, as the variable

Table 2 Similar issues drive donor and non-donor ideology

	Donor CFscores	Donor self-placements	Non-donor self-placements
Donor CFscores	1.00		
Donor self-placements	0.95	1.00	
Non-donor self-placements	0.83	0.93	1.00

Cell entries are Pearson's correlation coefficients between variable importance vectors from random forests using 25 CCES policy questions to explain each measure of ideology. All significant at $p < 0.05$

importance vectors for the forests that use self-placement as an outcome variable correlate at $r=0.93$. But the relationship is still quite high, at $r=0.83$, when I compare the vectors for models that predicted non-donor self-placements and donor CFscores.

Based on this evidence, it appears that despite their demographic differences, donors and non-donors draw on similar issues to establish their left–right ideology orientation. This allows us to engage the question of whether issue positions can be used as bridge data to impute CFscores for non-donors.

Using Issue Positions to Predict CFscores for Donors

The proposition that we can impute CFscores for non-donors, based on how their issue positions compare to those of donors, depends on the ability of those positions to amply predict CFscores. And notably, this has been the subject of debate. Hill and Huber (2017), in their analysis, used factor scaling to create a left–right policy score from the policy-based questions in the CCES. They found that while CFscores clearly distinguish donors with left versus right policy orientations, they were less able to explain variation among Democratic and Republican donors. Within these groups, correlations between CFscores and policy scores were $r=0.22$ for Democrats and $r=0.42$ for Republicans. In his response, Bonica (2019) regressed CFscores on each of the policy questions in the 2012 CCES, showing that 24 of the 25 questions significantly explain within-party variation in CFscores.

My analysis finds support for both narratives, but for the purposes of imputing CFscores for non-donors, I consider issue positions to be suitable bridge data. This is because they perform well across the full sample, and although they are weaker at the extremes of the ideological scale, they separate left from right, and are highly predictive across the middle of the observed range of CFscores (where most non-donors will be located).

To study the ability of issue positions to predict CFscores when using random forest machine learning, I subset my data into two groups: the “training” and the “testing” data. Random forest algorithms are sophisticated enough that, if the data are not subset like this, then their predictions are likely to be very close to the real values because of overfitting. As such, I use a random forest to train responses to the 25 CCES policy views onto the CFscores of 50% of donors in the dataset. These are the “training” data.³ Then, I use that model to predict the CFscores for the remaining

³ In terms of model specifications, the random forest is composed of 1,000 regression trees. Two covariates were randomly selected at each node, which split on the covariate that explained most variation in CFscores, until reaching a terminal node of no more than 10 observations.

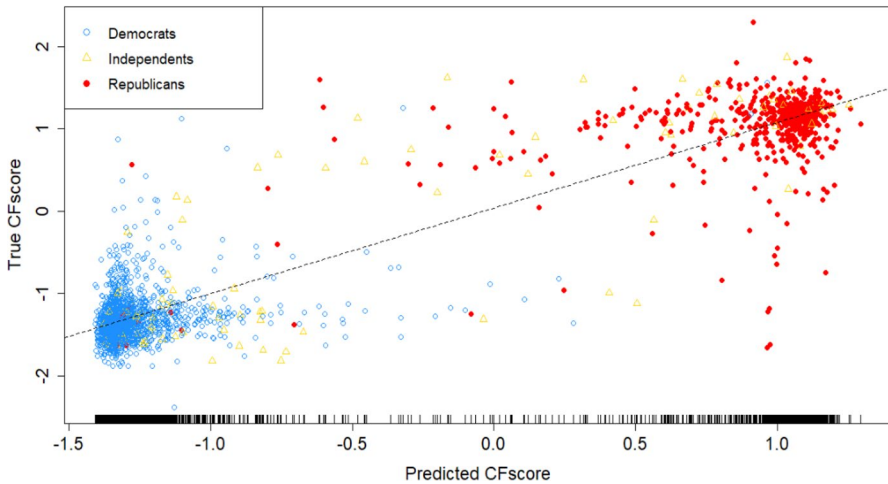


Fig. 2 Random forest predictions of donor CFscores ($r=0.94$)

50% (the “testing” data),⁴ and present the relationship between true and predicted values for this subset of donors.

Figure 2 presents the relationship between true and predicted values for the “testing” data. Each point represents an individual donor and is color-coded to reflect their partisanship, as reflected in the legend in the upper-right.⁵ The best-fit line for the relationship between predicted and CFscores is plotted by the dashed black line, and the rug plot along x-axis reflects the distribution of predicted scores.

The most important feature of this plot is seen in the large clusters of Democratic donors on the left and Republican donors on the right. The y-axis indicates that most Democrats held CFscores of between -1.0 and -1.5 , and most Republicans held scores close to 1.0 . The random forest model placed almost all of these donors in the correct location on the x-axis, yielding a high overall relationship of $r=0.94$ between true and predicted CFscores. In line with Hill and Huber’s (2017) findings, the model is more limited in its ability to predict within parties, with values correlating at $r=0.36$ for Democrats and $r=0.49$ for Republicans. This is as expected, as donors contribute for reasons besides ideology (e.g., to support local candidates or help their party in close races) and the ± 0.10 random noise added to the data to preserve anonymity makes it hard to explain small differences within the dense

⁴ Much like an OLS can be used to create fitted values for out-of-sample data, a random forest model trained on one set of data can be used to create predictions of the same outcome variable on another set, so long as the latter contains all the same covariates. Here, donors in both training and testing sets responded to the same battery of CCES questions. This allows me to run the testing responses through the trees of the random forest. At the end of each tree, when observations arrive at a terminal node, they are assigned that node’s mean value (from data in the original forest) as the tree’s estimate. As with the original forest, this produces hundreds of estimates for each observation, which are averaged together to create the forest’s final predictions.

⁵ “Leaners” who do not report an affiliation, but say they feel closer to one party or another are coded as belonging to that party.

left and right clusters that most donors fall into. However, for observations that fall between -1 and 1 , the relationship between true and predicted values is again high at $r=0.82$, and this is the range into which most of my imputed values will fall.

From this exercise, it is clear that random forest machine learning can use issue positions to distinguish between left and right and the points in-between. This gives me confidence that they are suitable as bridge data for imputing non-donor CFscores.

Imputing, Validating, and Evaluating CFscores for Non-donors

In the previous sections, I showed that issue positions can be used to predict CFscores, and that the “mapping” process by which issue positions translate to left–right ideology is similar for donors and non-donors. Establishing these allows me to impute CFscores for CCES respondents who did not donate to a political candidate, based on how their issue positions compare to those who did.

To begin, I run another random forest model that has the same technical specifications, but includes all 4432 donors as observations and also incorporates variables for age, income, and political knowledge. I include all donors here to bring as much information to the imputation of non-donor CFscores as possible, and I am no longer concerned about predicting donor scores out-of-sample. The age, income, and knowledge variables are included to model how these key differences between donors and non-donors may condition the way issue positions relate to left–right ideology.

The results of this imputation are summarized in Table 3, where the imputed CFscores (for non-donors) and true CFscores (for donors) is presented at different points of each party’s ideological distribution. As expected, most Democrats are placed to the left of most Republicans, and independents are primarily concentrated in the ideological middle. Further, even though most donors occupy the extreme left and right of the ideological scale, the random forest model still placed the majority of non-donors in the middle. Roughly half of Democrats and three-quarters of Republicans were assigned scores between -1 and 1 . These observations lend face validity to the imputed CFscores.

While Table 3 shows that imputed CFscores look as they should, a separate question is whether they *behave* as they should: for instance, by correlating with

Table 3 CFscores by partisanship for donors and non-donors

	10th Percentile	25th Percentile	Median	75th Percentile	90th Percentile
“True” values for donors					
Democrats	− 1.59	− 1.49	− 1.35	− 1.22	− 1.03
Independents	− 1.58	− 1.41	− 1.06	1.12	1.40
Republicans	0.64	0.96	1.14	1.25	1.38
Imputed values for non-donors					
Democrats	− 1.32	− 1.27	− 1.09	− 0.71	− 0.27
Independents	− 1.18	− 0.89	− 0.40	0.22	0.90
Republicans	− 0.43	0.18	0.80	1.03	1.10

measures and predictors of ideology. To convergently validate these scores, I correlate them with respondents' self-placements on the CCES ideological scale, as well as the first principal component of their responses to the 25 policy questions on the survey. The relationship between the imputed CFscores and these alternative measures is $r=0.69$ and 0.92 , respectively. I also check whether CFscores are more extreme when respondents are more knowledgeable about politics. Indeed, the absolute value of a respondent's imputed CFscore correlates with their score on a two-question quiz about the majority parties in Congress at $r=0.35$.

Imputed CFscores appear valid as a measure of ideology, but more open is the question of how they compare to alternative measures. To test this proposition, I conduct a bootstrap analysis that compares imputed CFscores among non-donors with their ideological self-placements and the score derived from the first principal component of their responses to policy questions.

The bootstrap analysis randomly draws 500 responses from the CCES and runs logistic regressions of US House vote choice on each measure of ideology. From the regressions' fitted values, I calculate and record the percentage of observations that each model predicts incorrectly (its "error rate"), and do this for 500 bootstrapped samples. I then compare the average error rates achieved using each measure: the imputed CFscores, ideological self-placements, and the PCA-scaled policy views.⁶

The results of this exercise are presented in the plots in Fig. 3. Each row corresponds with a measure of ideology, and the grey points represent the error rates observed (as indicated by the x-axis) in each of the 500 bootstraps. The black bars, meanwhile, show the 95% confidence interval around the mean.

Across all measures used, ideology does a good job of predicting US House vote, but the imputed CFscores perform the best. They achieve a mean error rate of 14.2% across the 500 bootstrapped samples. By way of comparison, this represents relative improvements of 13% and 29% over the mean error rates that result from using PCA scaling and ideological self-placement, respectively.

The results in Fig. 3 suggest that imputed CFscores are more predictive than other measures of ideology. However, we may also be interested in the quality of their *comparisons*. Imputed CFscores put non-donors in an ideological common space with other actors, including fellow citizens who donated to candidates. If this common space yields high-quality comparisons, then non-donors should behave similarly to donors nearby them in that space. By contrast, if the common space is rough or approximate, we should expect to see less similarity in how donors and nearby non-donors behave.

We can assess the quality of CFscores' common space by comparing it to a space created by the "direct comparison" approach. Recall that the "direct comparison" approach involves finding topics on which different groups have commonly expressed an opinion. In expectation, this should yield more precise comparisons at the expense of scalability. Left-right measures like CFscores allow immediate

⁶ The bootstrap analysis is used to compare the strength of using each measure, as if they were used over many different analyses. Regressions run using the full dataset similarly show CFscores to be more predictive of House vote than other measures.

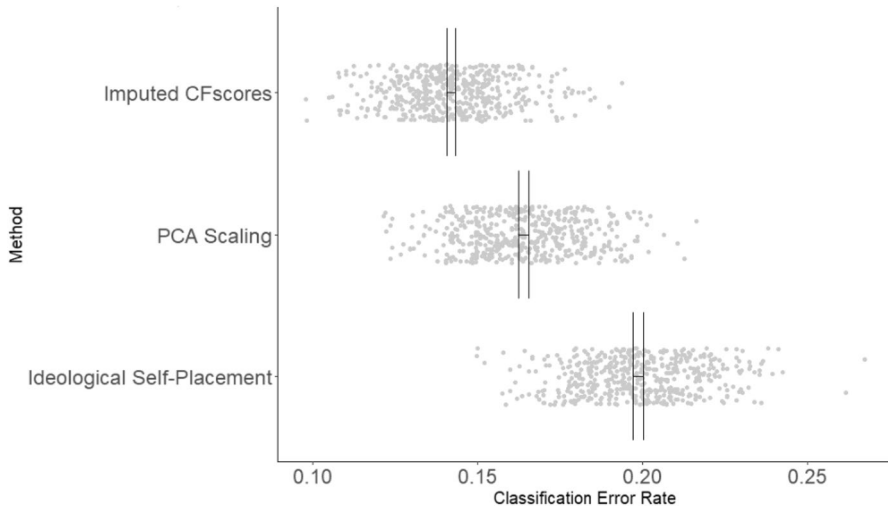


Fig. 3 Imputed CFscores predict US House vote with less error than other measures of ideology

comparisons between any actor they have been specified for, whereas “direct comparison” requires new data collection for each new analysis. But, these comparisons derive from exactly the same items in a way that summary left–right measures may not.

The fact that donors and non-donors responded to the same set of policy questions on the CCES allows us to set up a “direct comparison” between them. Following Jessee (2009) and Shor and Rogowski (2018), I reduce these policy questions to a unidimensional scale using principal component analysis. This sets up two common spaces by which to compare donors and non-donors. One is based solely on policy views, and the other (true and imputed CFscores) is based on donation patterns for donors, and a combination of donation patterns and policy views for non-donors.

To assess the quality of their comparisons, I use an algorithm to pair each donor with the non-donor whose policy views are most similar. Then, separately, I pair each donor with the non-donor whose imputed CFscore is closest to their real one. Finally, for each of these paired datasets, I calculate the frequency with which the donor and their paired non-donor cast the same presidential and US House ballots, and identify with the same political party.

The results of those comparisons are presented in the bottom two rows of Table 4. The first of these shows the rate of behavioral agreement when we pair donors with non-donors on their issue positions alone. The row below it shows the same rates, but when we compare donors and non-donors on their CFscores instead.

Across all three criteria measures, the comparisons via CFscores see more agreement, although the margins are very small.⁷ This is an informative result, because

⁷ Notably, we see this result despite the algorithm finding, on average, nearer pairings based on donor policy views than CFscores.

Table 4 CFscores and policy views yield comparisons of similar quality

Point of comparison	Agreement of donors and nearest non-donors on...		
	Presidential vote (%)	US House vote (%)	Party identification (%)
Policy views	95.5	90.8	83.9
True and imputed CFscores	95.7	92.7	85.9

we generally expect the effort that goes into finding direct points of comparison to yield a higher-quality common space than the comparison of ideal point estimates. However, beyond their predictive ability, CFscores also appear to function well in terms of setting up common space comparisons. In the next section, I further showcase their strength in setting up comparisons across the political system, with a substantive application that compares the effect of ideology on vote choice at different levels of government.

Examining the Role of Ideology in US Gubernatorial Elections

Governors are distinctive actors in the US political system. As executives of the 50 states, they heavily influence areas as diverse as bureaucratic administration (Haas & Wright, 1989), fiscal policy (Breunig & Koski, 2009), and the legislative agenda (Beyle, 1999), and they each operate in their own unique political contexts. Given the influence and distinctiveness of the role, it is perhaps unsurprising that gubernatorial elections often diverge from patterns set by races at the federal level. Despite the increasing nationalization of politics, fully 30 percent of governors in the 2010s were affiliated with the party that lost their state in the most recent presidential election, and that number exceeded 50 percent as recently as the 1980s (Sievert & McKee, 2019).

In the past, explanations for the preponderance of minority-party governors have focused on economic voting and ideology. The former of these refers to the premise that voters cast their ballots to reward (or punish) the incumbent party for a good (or bad) economy. Since governors are the highest-profile leader at the state level, several studies suggested that voters assess them and their party based on local conditions, separately from how the nation is doing as a whole (Atkeson & Partin, 1995; Stein, 1990; Svoboda, 1995). Meanwhile, other research focused on ideology and worked from the assumption that candidates should converge on the center to maximize their vote share (Downs, 1957). In their seminal study of state politics, Erikson et al. (1993) demonstrated that although party activists tend to align ideologically with the national party, they nominate candidates for governor who reflect the electorate's ideology—with Republicans nominating moderates in blue states, and Democrats in red. The importance of ideology in these elections was highlighted by Cook et al. (1994),

who showed that issue agreement with candidates on abortion, death penalty, and gun control rivaled the effect of economic retrospection in gubernatorial vote choice.

Recent years have brought renewed attention to this discussion, propelled by the greater availability of data on state-level conditions and public opinion. While studies have qualified the role of economic retrospection in light of partisan-motivated reasoning and attribution error (e.g., Brown, 2010), the importance of ideology appears clearer now than ever. Wolak and Parinandi (2022) show that gubernatorial approval is shaped in large part by how policy outcomes reflect the ideological preferences of state electorates, and Caughey and Warshaw (2019) find that candidates for governor face a stronger penalty in terms of vote share for ideological extremism than do candidates for US House or Senate. Warner (2022), in a validation of CFscores at the gubernatorial level, further shows that ideological moderates are elected to governor's mansions more often than they are to Congress.

Taken together, these studies suggest that ideology has more impact on vote choice in gubernatorial elections than it does in others. However, this has not been tested directly, as doing so would require scholars to place voters on the same ideological scale as candidates for multiple types of office. Using the “direct comparison” approach, it would take a herculean effort to find topics that candidates for governor and Congress—across many different states—had commonly registered opinions on, and survey data that asks voters about their candidate preferences for those offices and issue positions on the same the topics.

CFscores make such common-scale comparisons accessible, and by imputing scores for citizens, we can add voters to the large universe of actors for which these measures exist. In the analysis below, I use CFscores to calculate CCES respondents' relative ideological distance from the major-party candidates in gubernatorial and US House races held in 12 states on Election Day 2012. If ideological proximity is shown to matter more for gubernatorial vote choice, then this would give greater clarity as to why moderate governors are so commonplace. By contrast, if the effect of ideological proximity on vote choice is the same across gubernatorial and US House races, then this would suggest that it is something else about the nomination of extremist candidates—such as poor candidate recruitment, or higher turnout among the opposition—that more directly explains their relative underperformance in executive (as opposed to legislative) elections.

Research Design

To compare the effect of ideological proximity to the candidates on vote choice, I subset the 2012 CCES to respondents who lived in a state with a gubernatorial election that year and cast a ballot for governor and/or US House. Together, they lived in 12 states and 60 US House districts. I then specified logistic regressions to study vote choice in each type of race. The dependent variable in these regressions was whether (1) or not (0) a ballot was cast for the Republican candidate.

The key independent variable in each model is the relative ideological advantage of the Republican candidate. Creating this measure requires three ideological

points—one for both major-party candidates and one for the voter—and represents a methodological step up from the large number of studies that use only one or two (see Tausanovitch and Warshaw 2018 for an overview). By using estimates for both candidates and for the voter, we can observe not only how voters become more likely to support one candidate when they are ideologically nearer, but also when they are more distant from the opposition.

I calculate the Republican's ideological advantage using the equation in Shor and Rogowski's (2018) study of US House races.

$$\text{GOP ideological advantage} = |x_i - x_D| - |x_i - x_R|$$

The term x_i refers to the CFscore of voter i . For this term, the imputed values are used for non-donors, and true values are used for citizens who donated to a candidate. The x_D term refers to the CFscore of the Democratic candidate, and likewise, the x_R term refers to the Republican candidate's CFscore.

CFscores range from -2 to 2 , with larger values being more conservative. In the equation above, as the voter gets closer to the Republican, the resulting value gets larger. For example, a voter who is very conservative may hold the same CFscore as the GOP candidate, at 1.50 , while the liberal Democrat has a CFscore of -1.50 . This would result in a positive value for the GOP ideological advantage, as $|1.50 - (-1.50)| - |1.50 - 1.50| = 3$. Meanwhile, a value of zero would indicate that the voter is at the exact midpoint between candidates, and a negative value would indicate closer relative proximity to the Democrat.

To isolate the effect of ideology on vote, I also include several control variables. I account for campaign effects by taking the difference between the Republican and Democratic candidates' fundraising totals in millions of dollars, and presenting it as the GOP fundraising advantage (where negative values indicate that the Democrat raised more). At the individual level, I include partisanship as an identity-based factor that works separately from ideology to affect citizens' preferences at the ballot box. I also include race, age, sex, and education as demographic control variables that may affect vote choice. Finally, I include state and district fixed effects for the gubernatorial and US House vote models, respectively, to control for the contextual effects that are unique to each race.

Results and Discussion

The results of the two logistic regressions, on vote choice for governor and US House respectively, are presented in the columns of Table 5. In the first column, covariates are presented, and in the next two columns are the results of my model of gubernatorial vote choice. Then, the last two columns present the results of the US House model.

The coefficients in the tables have been exponentiated (from their original, logged-odds form) to create odds ratios that describe how a one-unit increase in each covariate multiplies the odds of voting Republican. The first row shows, for example, that in a gubernatorial election, a one-unit increase in the Republican's ideological advantage multiplies the odds of a voter selecting them by 2.60, or put

Table 5 Ideology equally affects vote choice in races for governor and US House

	Governor		US House	
	Odds Ratio	Conf. interval	Odds ratio	Conf. interval
GOP ideological advantage	2.60	[2.36, 2.84]	2.44	[2.21, 2.67]
GOP fundraising advantage	2.53	[1.21, 3.85]	0.49	[0.00, 1.28]
Democrat	0.13	[0.09, 0.16]	0.16	[0.11, 0.21]
Republican	3.47	[2.54, 4.41]	5.13	[3.61, 6.64]
White	2.21	[1.59, 2.84]	2.22	[1.48, 2.97]
Female	1.78	[1.43, 2.13]	1.07	[0.84, 1.31]
4-Year degree	0.83	[0.65, 1.00]	0.95	[0.72, 1.18]
Year of Birth	1.00	[0.99, 1.01]	1.00	[0.99, 1.00]
Intercept	2.36	[0.00, 28.60]	4.87	[0.00, 68.44]
Observations	5174		4583	

Results are odds ratios from a logistic regression explaining whether (1) or not (0) a voter selected the GOP candidate. Governor model includes state fixed-effects, US House model district fixed-effects

differently, increases those odds by 160%. By contrast, an odds ratio below 1.00 would reduce the odds of a Republican vote. Next to the odds ratios are the 95-percent confidence intervals around them. If the interval overlaps with 1.00, then the result should not be considered statistically significant.

The key result from these models is presented in the first row. In gubernatorial elections, a one-unit increase in a voter's relative proximity to the Republican candidate increases their odds of voting for them by 160%. This is statistically indistinguishable from the effect in US House elections,⁸ where a one-unit increase in the Republican's ideological advantage associates with a 144% increase in the odds of voter supporting them.

This result is surprising, because between the higher profile of gubernatorial elections, and the greater frequency with which moderates are elected, I had expected that voters would be more responsive to the ideological positioning of the candidates. However, the results in Table 5 negate that expectation. From them, it appears that voters are about equally as sensitive to the ideological positioning of gubernatorial candidates as they are US House candidates, when a variety of individual and contextual factors are controlled for. Given the relatively small number of races analyzed—12 for governor, and 60 for US House—it may be reasonable to question whether these null results are an artifact of case selection. This would most obviously occur if the 2012 gubernatorial elections were unusually sleepy or uncompetitive, to the point that voters invested less in making ideologically-motivated decisions. But that does not appear to be the case. One-third of the races were decided by 5% or less (including in the

⁸ This comparison can be made using the z-test below, where β represents the odds-ratios and $SE\beta$ their standard errors (Clogg et al. 1995), which yields a value of 0.93 ($p=0.35$) when the results from these models are used.

$$Z = \frac{\beta_1 - \beta_2}{\sqrt{(SE\beta_1)^2 + (SE\beta_2)^2}}$$

larger states of Indiana and Washington), and on a per-race financial basis, the elections for governor saw about six times more fundraising than those for US House (with an average of \$12.2 to \$2.1 million). That said, the use of elections in 2012 constitutes a “hard” test for the relative effect of ideology in gubernatorial elections, because they would have co-occurred with that year’s presidential race. More states hold their elections for governor in midterm years, which has been shown to reduce the influence of national politics over these races. Data collection efforts that expand imputed CFscores beyond the 2012 CCES would allow a more thorough investigation.⁹

The clarity of the results in Table 5, however—with similar coefficients on the ideological measures in both models—encourages us to take them at face value. This leads to the second-order question of *why* we would see a greater penalty for extremism in gubernatorial versus US House election returns (Caughey & Warshaw, 2019) and more moderate governors than US House representatives (Warner, 2022) if voters are no more responsive to candidate ideology in gubernatorial races.

One dynamic that may separately result from candidate extremism involves turnout effects. Hall and Thompson (2018) use a regression-discontinuity design to show that when an extremist just barely wins their party’s nomination for US House, it has a positive, causal effect on turnout among supporters of the other party. Miller (2022) expands on this to show that extremists also reduce voter roll-off among the opposition, such that voters are less likely to leave down ballot races blank when their less-preferred party has nominated an extremist. Both analyses were conducted at the US House level, and with the higher profile of gubernatorial races, it seems plausible that these dynamics would be more pronounced in them. Separately, candidate extremism may be endogenous to other factors, such as challenger quality or party disunity, that may also lead to lower vote shares as the profile of an office increases (Bond et al., 1985; O’Brian, 2019).

Next Steps and Conclusions

In this paper, I demonstrated that CFscores can be extended as a measure of citizen ideology, placing citizens in an ideological common space with a large and diverse set of political actors. To do so, I used data from the 2012 CCES and Hill and Huber’s (2017) replication data, which paired donors in the study with their CFscores in Bonica’s (2014) Database on Money, Ideology, and Elections. Despite the systemic differences between donors and non-donors, I showed that the way issues map onto ideology for them is similar, and that issues can predict CFscores with relative accuracy.

Further, I demonstrated that beyond the advantage of being a common-space measure, imputed CFscores perform better than other measures of ideology in terms of predicting

⁹ Along similar lines, one might question the statistical power of the models, given that vote choice for governor and US House overlap so significantly; only 12% of respondents reported a split ticket between the races. However, among the control variables, we see differences in the effect of gender and candidate fundraising across races, and in reduced models, the effect of partisanship is shown to be stronger in US House vote whereas no differences emerge for the ideology covariate.

vote choice. I also provided an example of their substantive utility, by conducting an analysis of electoral behavior that places citizens on the same ideological scale as candidates for multiple offices at the same time. This allowed a more detailed look at how gubernatorial elections compared to others than was previously possible, and yielded the unexpected result that ideology may not impact votes for governor more than for US House. While this analysis would benefit from the incorporation of additional years of evidence, it showcases the level of detail that will be possible in models of electoral behavior and government responsiveness if CFscores can be extended to the public.

In terms of methodological next steps, the clearest goal is to expand the temporal availability of imputed CFscores for non-donors, as they are currently only available for the 2012 CCES. Three approaches stand out as the most plausible. The first involves pairing survey respondents to their entries in Bonica's DIME database. Hill and Huber (2017) commissioned YouGov to do this for the 2012 CCES, but this can be undertaken independently by any organization that administers its own surveys, such as the ANES. This approach likely needs to incorporate a time lag, as the DIME database is updated only every few years (version 3 is currently available and runs through 2018). A similar approach would involve inaugurating a survey of political donors with questions that collate with those of other major studies, such as the CCES or ANES. This could be done on a relatively low-cost basis, as FEC data can be merged with email addresses purchased from a vendor to allow for an Internet-based survey.¹⁰ Extending CFscores to non-donors using this approach again incorporates the time lag—the DIME database must be updated first—but it would serve the dual purpose of providing continuing, high-quality data on a set of citizens (donors) whose influence over modern-era politics appears considerable (Canes-Wrone and Gibson 2019, Kujala, 2020). For both approaches, the issues of sample size and the number of questions seems germane. Hill and Huber attained the CFscores of 4432 donors who answered 25 policy questions, but the ANES has much smaller sample sizes, and 25 questions represents a survey of intermediate or greater length, depending on their complexity. An analysis in the supplementary materials uses the CCES data to show that two-thirds of the information gain (relative to a unidimensional scaling of policy views) presented in Fig. 3 is achievable when the random forest is run using only 250 respondents and half of the questions used here.

A final way to extend CFscores to non-donors would borrow from the “direct comparison” approach of comparing citizens to elites, where researchers find topics on which both groups have commonly registered an opinion. Elite surveys like the Political Courage Test have been used to place voters and candidates on the same scale (e.g., Shor & Rogowski, 2018), and if their responses are predictive of CFscores, then the questions asked in common with citizen surveys may be used as bridge data to map citizens onto the common space. Researchers taking this approach would need to pay careful attention to non-response bias among

¹⁰ According to the FEC website, “In AO 1986–25, the Commission concluded that individual contributor information may be used for bona fide academic research projects that do not involve the sale or use of that information for a commercial purpose or for soliciting contributions.” Still, researchers wishing to conduct a donor survey in this manner should review their plans with their institution's IRB and general counsel.

candidates, and would need to separately validate whether the same issue positions that explain candidate CFscores explain the way citizens identify ideologically. However, to the latter concern, because candidate CFscores are derived from the patterns of donors, and because donors and non-donors appear to weigh issues similarly (see Table 2), this may not be as large a logical leap as it seems.

The discussion above, and this paper more broadly, highlight the fact that political science research is at an exciting point. Almost since the advent of quantitative research into politics, scholars have sought to compare citizens and elites directly, and while this is sometimes achieved on individual issues or for a particular set of actors, advances like Bonica's (2014) CFscores and Barberá's (2015) Twitter-based scores place large universes of political actors onto the same ideological scale.

Extending scores like these to the public enables comparisons at a larger scale than would have been considered possible even 10 years ago. In the analysis here, I made a one-to-one comparison of the role of ideology on vote choice across different types of elections, incorporating the ideal points of challengers, incumbents, and voters alike. Common-space scores can be used to make other such comparisons, to assess the role of ideology in elections that are less commonly studied, or to place citizens and their representatives on the same scale in models of government responsiveness. Ultimately, while the exercise here represents just one step in the process, it illustrates how close political science is to creating a "true" ideological common space—one that incorporates almost every actor in the political system—and speaks to the benefits in terms of measurement quality and analytic versatility that would come from the effort involved.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11109-023-09873-y>.

Acknowledgements This is a much-evolved version of my first project in graduate school, so thanks are due all around. Eric Plutzer reviewed multiple drafts of the manuscript. Suzanna Linn and Michael Berkman provided feedback when it was just a Master's thesis. Eric Manning and others at MPSA 2021 also offered valuable suggestions, and comments from the three anonymous reviewers helped this project to achieve its full scope. Any errors are mine alone.

References

- Afzal, A., Aabid, A., Khan, A., Khan, S. A., Rajak, U., Verma, T. N., & Kumar, R. (2020). Response surface analysis, clustering, and random forest regression of pressure in suddenly expanded high-speed aerodynamic flows. *Aerospace Science and Technology*, *107*, 106318.
- Aldrich, J. H., & McKelvey, R. D. (1977). A method of scaling with applications to the 1968 and 1972 presidential elections. *American Political Science Review*, *71*(1), 111–130.
- Atkeson, L. R., & Partin, R. W. (1995). Economic and referendum voting: A comparison of gubernatorial and senatorial elections. *American Political Science Review*, *89*(1), 99–107.
- Bailey, M. (2016). *Measuring ideology on the courts*. *Routledge Handbook of Judicial Behavior*. Routledge.
- Barberá, P. (2015). Birds of the same feather tweet together: Bayesian ideal point estimation using Twitter data. *Political Analysis*, *23*(1), 76–91.
- Berry, W. D., Fording, R. C., Ringquist, E. J., Hanson, R. L., & Klarner, C. E. (2010). Measuring citizen and government ideology in the US states: A re-appraisal. *State Politics & Policy Quarterly*, *10*(2), 117–135.
- Beyle, T. (1999). The governors. *Politics in the American states: A comparative analysis*, *7*, 191–231
- Bond, J. R., Covington, C., & Fleisher, R. (1985). Explaining challenger quality in congressional elections. *The Journal of Politics*, *47*(2), 510–529.

- Bonica, A. (2014). Mapping the ideological marketplace. *American Journal of Political Science*, 58(2), 367–386.
- Bonica, A. (2016). Avenues of influence: On the political expenditures of corporations and their directors and executives. *Business and Politics*, 18(4), 367–394.
- Bonica, A. (2018). Inferring roll-call scores from campaign contributions using supervised machine learning. *American Journal of Political Science*, 62(4), 830–848.
- Bonica, A. (2019). Are donation-based measures of ideology valid predictors of individual-level policy preferences? *The Journal of Politics*, 81(1), 327–333.
- Bonica, A., Chen, J., & Johnson, T. (2015a). Senate gate-keeping, presidential staffing of 'inferior offices', and the ideological composition of appointments to the public bureaucracy. *Quarterly Journal of Political Science*, 10(1), 5–40.
- Bonica, A., Chilton, A. S., Goldin, J., Rozema, K., & Sen, M. (2017). The political ideologies of law clerks. *American Law and Economics Review*, 19(1), 96–128.
- Bonica, A., Jowei, C., & Tim, J. (2015b). Senate gate-keeping, presidential staffing of "inferior offices," and the ideological composition of appointments to the public bureaucracy. *Quarterly Journal of Political Science*, 10(1), 5–40.
- Bonica, A., Rosenthal, H., Blackwood, K., & Rothman, D. J. (2020). Ideological sorting of physicians in both geography and the workplace. *Journal of Health Politics, Policy and Law*, 45(6), 1023–1057.
- Bonica, A., & Woodruff, M. J. (2015). A common-space measure of state supreme court ideology. *The Journal of Law, Economics, and Organization*, 31(3), 472–498.
- Breunig, C., & Koski, C. (2009). Punctuated budgets and governors' institutional powers. *American Politics Research*, 37(6), 1116–1138.
- Broockman, D., & Malhotra, N. (2020). What do partisan donors want? *Public Opinion Quarterly*, 84(1), 104–118.
- Brown, A. R. (2010). Are governors responsible for the state economy? Partisanship, blame, and divided federalism. *The Journal of Politics*, 72(3), 605–615.
- Brown, C. W., Jr., Hedges, R. B., & Powell, L. W. (1980). Modes of elite political participation: Contributors to the 1972 presidential candidates. *American Journal of Political Science*, 24, 259–290.
- Burke, R., Kirkland, J. H., & Slapin, J. B. (2021). Party competition, personal votes, and strategic disloyalty in the US States. *Political Research Quarterly*, 74(4), 1024–1036.
- Cahoon, L. S., Hinich, M. J., & Ordeshook, P. C. (1976). A Multidimensional Statistical Procedure for Spatial Analysis.
- Caughey, D., & Warshaw, C. (2019). Electoral Penalties for Ideological Extremism in US Federal and State Elections.
- Cook, E. A., Jelen, T. G., & Wilcox, C. (1994). Issue voting in gubernatorial elections: Abortion and post-Webster politics. *The Journal of Politics*, 56(1), 187–199.
- Cooper, C. A., Knotts, H. G., & Ragusa, J. (2016a). The constrained Governor: Exploring gubernatorial decision making on senate appointments. *Political Research Quarterly*, 69(3), 482–494.
- Downs, A. (1957). An economic theory of political action in a democracy. *Journal of Political Economy*, 65(2), 135–150.
- Enns, P. K. (2015). Relative policy support and coincidental representation. *Perspectives on Politics*, 13(4), 1053–1064.
- Enns, P. K., & Koch, J. (2013). Public opinion in the US States: 1956 to 2010. *State Politics & Policy Quarterly*, 13(3), 349–372.
- Erikson, R. S., Wright, G. C., Wright, G. C., & McIver, J. P. (1993). *Statehouse democracy: Public opinion and policy in the American states*. Cambridge University Press.
- Gilens, M., & Page, B. I. (2014). Testing theories of American politics: Elites, interest groups, and average citizens. *Perspectives on Politics*, 12(3), 564–581.
- Gooch, A., & Huber, G. A. (2020). how issue positions affect candidate performance: Experiments comparing campaign donors and the mass Public. *Political Behavior*, 42(2), 531–556.
- Granberg, D., & Brown, T. A. (1992). The perception of ideological distance. *Western Political Quarterly*, 45(3), 727–750.
- Grant, J. T., & Rudolph, T. J. (2002). To give or not to give: Modeling individuals' contribution decisions. *Political Behavior*, 24(1), 31–54.
- Haas, P. J., & Wright, D. S. (1989). Public policy and administrative turnover in state government: The role of the governor. 1. *Policy Studies Journal*, 17(4), 788–803.
- Hall, A. B., & Thompson, D. M. (2018). Who punishes extremist nominees? Candidate ideology and turning out the base in US elections. *American Political Science Review*, 112(3), 509–524.

- Hill, S. J., & Huber, G. A. (2017). Representativeness and motivations of the contemporary donorate: Results from merged survey and administrative records. *Political Behavior*, 39(1), 3–29.
- Jessee, S. A. (2009). Spatial voting in the 2004 presidential election. *American Political Science Review*, 103(1), 59–81.
- King, G., Murray, C. J., Salomon, J. A., & Tandon, A. (2004). Enhancing the validity and cross-cultural comparability of measurement in survey research. *American Political Science Review*, 98(1), 191–207.
- Kujala, J. (2020). Donors, primary elections, and polarization in the United States. *American Journal of Political Science*, 64(3), 587–602.
- Lenz, G. S. (2009). Learning and opinion change, not priming: Reconsidering the priming hypothesis. *American Journal of Political Science*, 53(4), 821–837.
- Lerner, J. Y. (2017). *Using Ideas As My Maps: Tracing the Flow of Ideas Through the Legislative Process* (Doctoral dissertation, Duke University).
- Lo, J., Proksch, S. O., & Gschwend, T. (2014). A common left-right scale for voters and parties in Europe. *Political Analysis*, 22(2), 205–223.
- Martin, G. J., & Peskowitz, Z. (2018). Agency problems in political campaigns: Media buying and consulting. *American Political Science Review*, 112(2), 231–248.
- McCarty, N., Poole, K., & Rosenthal, H. (2005). *Polarized America: the dance of ideology and unequal riches*. mit Press.
- Miller, M. G. (2022). Candidate extremism and voter roll-off in US house elections. *Legislative Studies Quarterly*, 47(4), 791–821.
- Miller, W. E., & Stokes, D. E. (1963). Constituency influence in Congress. *American Political Science Review*, 57(1), 45–56.
- O’Brian, N. A. (2019). One-party states and legislator extremism in the US House, 1876–2012. *The Journal of Politics*, 81(4), 1223–1239.
- Shor, B., & Rogowski, J. C. (2018). Ideology and the US congressional vote. *Political Science Research and Methods*, 6(2), 323–341.
- Sievert, J., & McKee, S. C. (2019). Nationalization in US Senate and gubernatorial elections. *American Politics Research*, 47(5), 1055–1080.
- Stein, R. M. (1990). Economic voting for governor and US senator: The electoral consequences of federalism. *The Journal of Politics*, 52(1), 29–53.
- Svoboda, C. J. (1995). Retrospective voting in gubernatorial elections: 1982 and 1986. *Political Research Quarterly*, 48(1), 135–150.
- Tausanovitch, C. (2019). Why are subnational governments responsive? *The Journal of Politics*, 81(1), 334–342.
- Tausanovitch, C., & Warshaw, C. (2014). Representation in municipal government. *American Political Science Review*, 108(3), 605–641.
- Tausanovitch, C., & Warshaw, C. (2018a). Does the ideological proximity between candidates and voters affect voting in US house elections? *Political Behavior*, 40(1), 223–245.
- Warner, S. (2022). Measuring executive ideology and its influence. *State Politics & Policy Quarterly*. <https://doi.org/10.1017/spq.2021.34>
- Wolak, J., & Parinandi, S. (2022). Does the public hold governors accountable for policy outcomes? *Political Research Quarterly*, 75(4), 1051–1064.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.