Political Science Research and Methods Vol 4, No. 1, 97–121 January 2016
© The European Political Science Association, 2015 doi:10.1017/psrm.2015.5

Measuring State and District Ideology with Spatial Realignment

JAMES E. MONOGAN III AND JEFF GILL*

We develop a new approach for modeling public sentiment by micro-level geographic region based on Bayesian hierarchical spatial modeling. Recent production of detailed geospatial political data means that modeling and measurement lag behind available information. The output of the models gives not only nuanced regional differences and relationships between states, but more robust state-level aggregations that update past research on measuring constituency opinion. We rely here on the spatial relationships among observations and units of measurement in order to extract measurements of ideology as geographically narrow as measured covariates. We present an application in which we measure state and district ideology in the United States in 2008.

Political Scientists face substantial data limitations when measuring public opinion. Much of the research on representation and the role of the public in politics requires good measures of the constituencies to which elected officials are accountable. In the American context, these constituencies may be the residents of a state (who elect their governor and senators), a congressional district (electing their member of the House of Representatives), a state legislative district (electing their member to the state lawmaking body), or any number of other geographically referenced areas that choose a policymaker of consequence. Survey data that provide an estimate of constituency ideology that is not only reliable, but comparable with other districts, are hard to come by. Hence, the discipline of Political Science has had to develop a variety of disparate strategies to address this scarcity of data.

In this methodological work, we develop a new approach for modeling public sentiment by micro-level geographic region based on Bayesian hierarchical spatial modeling using multiple geocoded data sources. Unfortunately, the current state of geographically oriented measurement and modeling in Political Science currently lags behind the stunning amount of raw geospatial political data available for analysis. We seek to resolve this discrepancy with model output that produces nuanced regional relationships and contrasts between constituencies at varying levels, as well as robust district-level aggregations. By exploiting the spatial relationships among observations and units of measurement, we extract measurements of ideology as geographically narrow as available covariate measures. Our approach broadens the process of areal (defined regions) spatial modeling

* James E. Monogan III, Assistant Professor, Department of Political Science, University of Georgia, Athens, GA 30602 (monogan@uga.edu). Jeff Gill, Professor, Departments of Political Science, Biostatistics and Surgery, Washington University, One Brookings Drive, Seigle Hall, St. Louis, MO 63130-4899 (jgill@wustl.edu). The authors thank Robert S. Erikson, Bradley P. Carlin, Guy D. Whitten, Stephen Ansolabehere, Robert J. Franzese, Jude Hays, Paulo J. Ribeiro Jr., and the anonymous reviewers for their helpful assistance. This study was supported in part by resources and technical expertise from the Georgia Advanced Computing Resource Center, a partnership between the University of Georgia's Office of the Vice President for Research and Office of the Vice President for Information Technology. Questions may be directed to James Monogan as corresponding author. Complete replication information and our estimates of ideology in 2008 are available at our Dataverse page: http://hdl.handle.net/1902.1/22006. To view supplementary material for this article, please visit http://dx.doi.org/10.1017/psrm.2015.5

https://doi.org/10.1017/psrm.2015.5 Published online by Cambridge University Press

in Political Science by producing a distributed product that allows other researchers to produce aggregation at all practical political levels in American politics: states, congressional districts, census tracts, MSAs, state legislative districts, cities, and counties. This is a new paradigm for spatial political analysis that is flexible on aggregation and draws important information from social, economic, and political information at the most local level.

The end product of this work is a model-based smooth public opinion measure based on the appropriate treatment of hierarchical covariates. To obtain this result, we identify discrete non-nested hierarchies in a collection of data sets and smooth them with *Bayesian kriging* to produce a model in the form of a *posterior density blanket* that draws spatial estimates of ideology and other public sentiment from both smoothed geographic terms as well as principled covariates across those levels. By *posterior density blanket*, we mean a smoothed three-dimensional surface over the area of interest that results from the hierarchical specification along with the Bayesian kriging step, and is smoothed such that the variability of the predictions is less than the variability of the original observations. This density blanket is similar to the *spatial relative risk function* long employed in spatial epidemiology (Lawson and Williams 1993; Kellsall and Diggle 1995). The process described accounts for gradual spatial effects rather than arbitrarily sharp effects. This smoothing approach has not been done in the social sciences, even though there are obvious examples where boundaries are not socially determined.

We also generate freely available software that will enable other users to apply any division of interest and obtain smooth measures of latent ideology for their research purposes. Furthermore, the Bayesian hierarchical spatial models developed here can be used for any spatial or areal data analysis process where prediction is important. These models can be customizable beyond geographic measures of public opinion in the United States, given additional relevant data. For this project, the output includes our estimates of state-level ideology, and we also create estimates of congressional districts in Virginia to illustrate the capabilities of our method.

This paper proceeds first by describing the background theory of Bayesian hierarchical spatial modeling. Next, we describe the kind of model we build in order to forecast constituency ideology. Third, we describe the model we use in our current application, given current software limitations. Fourth, we present the forecasts our model produces for 2008 data in the lower 48 states plus the District of Columbia, and in congressional districts in Virginia, along with analyses that show how our measure compares with a simple measure based on survey subunits and multi-level regression with post-stratification. Finally, we discuss the implications of our results and the process for building a more computationally efficient and theoretically robust forecasting algorithm.

BACKGROUND

Studies in Political Science frequently fail to measure accurately the attributes of areal data units such as congressional districts. This usually means measuring an attribute of a district by weighting measures of component counties by the percentage of their area or weight in the district or, more crudely, simply using the measured value of a district's largest county while ignoring the fact that district lines are often drawn strategically. For example, if data on average income, voter turnout, religiosity, or any other factor were available at the county level, but not the district level, then a researcher would face this dilemma. Worse yet, political and social boundaries are often *non-nested*, meaning that there is no systematic hierarchy of geographic units laid out by governments. For instance, congressional districts can cut through counties and cities, census tracts do not tessellate counties, and other non-nested structures exist. This means that researchers are forced to make difficult choices about areal units of analysis in their studies.

https://doi.org/10.1017/psrm.2015.5 Published online by Cambridge University Press

In research using American congressional districts as the unit of analysis, the most pervasive measure of electoral ideology is the president's share of the two-party vote in the district. In principle, the presidential candidates are the same nationwide, so this measure will vary with the location of the median voter in a district. As Kernell (2009) demonstrates, however, ideologically *ranking* states or districts using presidential vote shares is only possible when the dispersion of voters is similar, and *rating* constituencies on a common cardinal scale is impossible with electoral data. Further, voting behavior may be based on a variety of non-ideological concerns that may not be constant across region, thereby inducing measurement error.

In state politics research, a wider variety of measurement techniques has emerged (for a more extensive review of these methods, see Cohen 2006, 6–10). One of the most pervasive techniques is the seminal measurement of state-level ideology by Erikson, Wright and McIver (1993), which pools several *CBS News*/*New York Times* polls together to obtain sufficiently large samples for each state. Pooling enough surveys over time requires the assumption that public ideology remains static, an assumption later maintained in several other measures of public sentiment (Lieberman and Shaw 2000; Norrander 2001; Brace et al. 2002; Park, Gelman and Bafumi 2004; Park, Gelman and Bafumi 2006; Lax and Phillips 2009). As an alternative perspective, Berry et al. (1998) later made the case that state ideology varies dynamically with a measure based on interest group ratings of congressional candidates, and this alternate view also appears in a wide array of studies (Gray 1976; Jacoby and Schneider 2001; Beyle, Niemi and Sigelman 2002; Brace et al. 2004; Erikson, Wright and McIver 2006).

Note, however, that model-based means of measurement, such as the post-stratification technique used by Park, Gelman and Bafumi (2004), Park, Gelman and Bafumi (2006), and Lax and Phillips (2009) are more amenable to dynamic change than many other techniques for static measurement. In fact, Pacheco (2011) illustrates how useful the post-stratification method can be for measuring time-variant state-level opinion. In this paper, we develop a new model-based technique that, similarly, can be used to generate a single static snapshot of opinion or can be fitted over several replicate surveys to create a dynamic measure, as the researcher sees fit. In this way, we develop a measurement strategy that can deal with data limitations that become more severe the smaller the constituency. Further, prior measures do not offer much attention to regional heterogeneity within states, so our technique offers an effective way to spatially model the effects of these discrepancies.

Strategy

The answer to this problem is to create a model-based smooth *density blanket* for the concept of interest (primarily ideology here) such that arbitrary boundaries are immaterial. We do this with a Bayesian hierarchical spatial model that uses kriging to analyze geocoded survey responses, and then we use this analysis to forecast areal units. A kriging estimate uses an optimized linear combination of the data at observed points to fill-in unobserved points, conditional on the spatial configuration of the observed data and spatial separation between the observed and estimated points (Tam Cho and Gimpel 2007; Cressie and Wikle 2011), thus producing a smooth density blanket over the complete geographic area of interest with lower variance than the original data. Once this density blanket exists, as determined from a posterior distribution of the chosen outcome and the subsequent smoothing, we and others can create any geographic division that a research question requires by slicing vertically such that divisions remain smooth and separated.

The idea of a smoothed density blanket is illustrated in Figure 1. The rectangular regions on the plane defined by the North–South/East–West dimensions indicate areal units of interest,

https://doi.org/10.1017/psrm.2015.5 Published online by Cambridge University Press

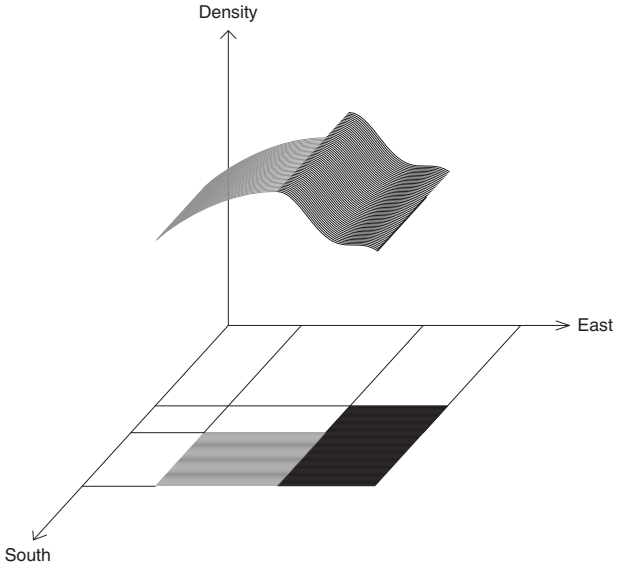


Fig. 1. Smooth surface over two non-overlapping contiguous geographic regions

which could be states, counties, etc. Two regions are highlighted in gray and black for illustration, although in our real analysis the entire United States is analyzed. Now above these two regions is the smoothed density blanket corresponding to the areal boundaries below. Notice that they are smooth at their boundary. Thus, for any region that analysts care about there is a smoothed function in the density dimension that seamlessly joints with that of its neighbors. As an analogy, consider the density blanket to be a sheet of cookie dough. Our algorithms can provide a cookie cutter in any shape that the cook wants. Of course, common "slices" will be by state, county, congressional district, or census tract. In this way, we provide a more realistic approach to modeling spatial variance for political effects.

How Does Our Approach Compare With Conventional Methods?

Three approaches to measuring local-level ideology are particularly pervasive in the literature: subsetting a survey, using election returns, and post-stratification. In the subsetting approach, the researcher simply subsets a large survey by a geographic distinction of interest and aggregates the subsets of data (e.g., Erikson, Wright and McIver 1993). There are a couple of problems with this approach. First, the sample size can become quite small in some cases, perhaps to the point of not observing any data for some units (e.g., North Dakota and Wyoming may be states with few or no respondents in a national survey). The second issue is that survey data often are not designed to be representative of a subsample of interest. The American National Election Study, for example, stratifies on region and is therefore not necessarily representative of each state. Hence, state subsamples can be misleading.

A second approach for measuring public sentiment is to use election returns as a proxy for district ideology. The easiest way to examine this is the share of the presidential vote in a district (Erikson and Wright 1980; Ansolabehere, Snyder and Stewart 2001). The assumption in this case is that because the candidates' ideologies are constant nationwide, the vote share will change only in response to the median voter. Alternatively, Berry et al. (1998) create a voting-based measure

https://doi.org/10.1017/psrm.2015.5 Published online by Cambridge University Press

that uses votes in congressional races, but measures the ideology of both incumbents and challengers to place each district on the same scale. Although vote-based measures use abundant data that are simple to gather, there are concerns with this kind of measure as well. First, vote choice is conceptually distinct from ideology. How an individual behaves at the ballot box may not be based on general ideology, but also regional appeals, personality traits, or economic well-being, thereby inducing added measurement error. Second, the vote share in a district is based solely on who turned out to vote and does not capture the views of non-voters. Finally, vote choice alone may be a misleading measure in that it does not account for the relative dispersion of ideological positions in a district (Kernell 2009).

A third approach seeks to resolve the data limitations and ideological dispersion issues present with the subsetting and proxy methods. *Post-stratification* models individual responses to survey questions. Then, responses are predicted for every combination of covariates and weighted based on known population quantities for the region of interest (Park, Gelman and Bafumi 2004; Park, Gelman and Bafumi 2006; Tausanovitch and Warshaw 2013). This method draws from weighting techniques also described in Gelman and Little (1997). Earlier work with such weighting was done by Pool, Abelson and Popkin (1965), Weber and Shaffer (1972), and Weber et al. (1972). Jackson (1989), seeking to address the constituency measurement problem, developed a technique wherein a model is fit to national-level survey data and then used to forecast public opinion in small area constituencies using average values of predictors in the constituencies.[1] Adding these weighted predictions presents an estimate of public ideology that also includes a measure of uncertainty. By resolving many of the problems posed by other measurement strategies, post-stratification presents one of the best means of capturing public ideology by region.

However, we believe that the principle of post-stratification can be further improved upon because no current method of public opinion measurement effectively incorporates the information geography offers. Post-stratification offers a structural model for what shapes individuals' ideological and issue preferences and the use of such a model to improve estimates of aggregate opinion. Besides observable and measured predictors that shape opinion preferences, though, there may be unobservable factors that shape individuals' preferences as well. These may be factors such as local discourse, variation in the meaning of partisan labels in local elections, commitment to certain values, membership in labor unions, or any number of other potential factors. Normally, unobservable input variables are relegated to the disturbance term in a model. However, most models assume that each respondent's disturbance is independent of all other respondents' disturbances. This assumption is not fair to make.

For a theoretical background on why we expect individuals who are geographically proximate to share similar political attitudes, Gimpel and Schuknecht (2003, 2–4) describe two different approaches to understanding regionalism in politics—a compositional approach and a contextual approach. First, the compositional approach maintains that political behavior bears similarity within a region because economic interests, racial origin, ethnic ancestry, religion, social structure, and other factors tend to be similar among a region's residents (Gastil 1975; Garreau 1981; Fischer 1989; Lieske 1993). In principle, under the compositional approach, if all of these factors could be included in an empirical model, then individuals' behavior would not show added similarities once the variables have been accounted for. However, it is often impossible to measure every relevant demographic and socioeconomic variable, or even to identify every relevant input in sufficient detail to include it in a model. (For example, would

[1] This technique also is used in applications for understanding representation presented in Jackson and King (1989) and Jackson (2008).

https://doi.org/10.1017/psrm.2015.5 Published online by Cambridge University Press

it be sufficient to know the relative importance of farming and manufacturing in an area, or are unique attributes of different crops and different products also important to regional culture?) As some relevant inputs may be overlooked or unmeasured, it is safer to assume that neighboring individuals will have a relatively similar political outlook, even holding the predictors constant.

The second approach to understanding regionalism, the contextual approach, offers the idea that citizens' political attitudes and behaviors are influenced by political socialization and by interactions with other citizens in their social network. When citizens influence each other by discussing politics or sharing information, or when they are socialized with similar values, their opinions and actions will tend to be similar (Putnam 1966; Putnam 1993; Huckfeldt and Sprague 1995; DeLeon and Naff 2004; Djupe and Sokhey 2011). "The first place to look for political networks is within the immediate physical proximity of each individual" (Sinclair 2012, 26). Therefore, we expect under the contextual approach as well that geographically proximate individuals will have relatively similar opinions, even in a general model.

As empirical support for this point, Erikson, Wright and McIver propose that, "the unique political cultures of individual states exert an important influence on political attitudes" (1993, 48). This notion builds on past work by scholars such as Elazar (1966), who proposes that states can be categorized based on an individualist, moralist, or traditionalist view of government's role. Erikson, Wright and McIver (1993, 56–68) proceed to show that a higher proportion of variance in ideology and partisanship can be explained by state-level dummies than by demographic information, though state residuals also will pick up some of the effects of unmeasured individual-level variables. This result also holds in the study of urban areas, where DeLeon and Naff (2004, 703) conclude that political culture shapes the impact of identity on public opinion and political participation. It appears, then, that citizens in nearby locales do share a similar political outlook.

Our method of kriging increases the ability to reasonably capture the effects of political culture, omitted predictors, and social context by including weighted neighbors' residuals in forecasts of public opinion. In neighboring regions, residents are likely to share a similar political culture, similar values of unobserved predictors, and similar political socialization. For example, western Kentucky and southeast Illinois are similar places that are likely to be populated with similar people, both in cultural and demographic terms. However, none of the three aforementioned approaches explicitly attempts to capture this in the measurement strategy.

Provided that geographically proximate citizens do share similar political views, then kriging will improve upon current measurement strategies in three ways. First, the kriging approach improves estimates of the structural model of individual opinion relative to methods like post-stratification. By dropping the assumption that each individual's response is independent of all others, we allow the disturbances in individuals' responses to survey questions to covary conditionally on how geographically proximate they are. Doing this yields more efficient parameter estimates than the wrongful assumption that the disturbances are independent. Hence, the predictive structural model ought to be more accurate.

Second, in the process of allowing spatial autocorrelation, we estimate a surface of what the spatial random effects are (see again Figure 1). By fitting this kriging surface and then using the values of this surface to predict ideology or issue opinion in a region, the added information will further raise the quality of our predictions. To the degree that unmeasured variables and political culture shape public opinion, the crafting of this surface ought to better incorporate these features into the forecast. Regardless of how small the constituency's area is, a reasonable projection of these factors is incorporated via kriging. As the kriging surface is identified

https://doi.org/10.1017/psrm.2015.5 Published online by Cambridge University Press

through the spatial location of each respondent, fitting this surface allows us to separate the variance not explained in the mean model into variance that spatially clusters and variance that is truly stochastic. This further reduction of variance that cannot be used in prediction improves our measures.

Third, every forecast should report a level of uncertainty in prediction. In our case, this takes the form of credible surfaces around the kriged density blanket. Predictions for larger areas (such as entire states) will be more certain than predictions for smaller areas (such as congressional or state legislative districts), and so these credible surfaces on either side of the estimate will be smaller in the former case. Naturally, the same would hold for any other method, including survey disaggregation and post-stratification, but by including the information of the spatial surface in our model and predictions, we can reduce the error of our areal predictions to be as small as possible, whereas providing accurate measures of uncertainty at any level of aggregation.

MODEL OVERVIEW

The first stage is to specify a Bayesian hierarchical model that accounts for aggregation in the data that we will use: individuals (in our case, survey respondents) are nested in regions, and the regions have associated covariates. These multi-level approaches make sure that the levels in the data are treated appropriately in the model and we do not assign regional or nesting variables directly to individuals (Gelman and Hill 2007; Gill and Womack 2012).

Basic Spatial Specification

Consider $\mathbf{X}^*$, an $n \times k$ positive definite matrix of explanatory variables organized down columns with a leading column of ones. Denote for the ith case, $\mathbf{x}_{ij}$ is a vector of explanatory variable values, and the j index is a reminder of nesting in the jth regional group. Furthermore, partition the $n \times k$ matrix $\mathbf{X}^*$ into two submatrices that track the nesting distinction:

$$\underset{n\times k}{\mathbf{X}^*} = \left[\begin{array}{c|c} \underset{n\times p}{\mathbf{X}} & \underset{n\times q}{\mathbf{Z}} \end{array}\right], \tag{1}$$

where $\mathbf{X}$ is the matrix of variables that will have *unmodeled* coefficients (receiving standard point estimates) and $\mathbf{Z}$ the matrix of explanatory variables that will have *modeled* coefficients through a hierarchy and therefore will be described distributionally. Finally, specify an error structure, ϵ, containing ϵ_{ij} values that are all independent and identically distributed according to $\epsilon_{ij} \sim \mathcal{N}_n(0, \sigma^2)$, $\sigma^2 > 0$. The complete model for our outcome is now the simple form:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \epsilon, \tag{2}$$

where $\boldsymbol{\beta}$ is a p-length vector of coefficient estimates and $\mathbf{u}$ the q-length vector of *random effects* with the assumption of multivariate normality with mean zero and positive definite variance–covariance matrix, such that $\mathbf{u} \sim \mathcal{N}_q(0, \Delta)$. This produces a method of accounting for structured data through regression equations at different hierarchical levels in the data. These higher-level models are describing *distributions* at the level just beneath them for the coefficient that they model as if it were itself an outcome variable. In this way, multi-level models are highly symbiotic with Bayesian specifications because the focus in both settings is on making supportable distributional assumptions.

This hierarchical approach models the systematic component, but does not yet incorporate the smoothing that we require for our purposes. The latter is done with kriging through the variance

https://doi.org/10.1017/psrm.2015.5 Published online by Cambridge University Press

TABLE 1 *Four Candidate Parametric Semivariograms and their Relative Fit*

Model	Semivariogram Function	BIC
Exponential	$\tau^2+\sigma^2(1-\exp(-\phi d_{ij}))$	271,018
Matérn ($\nu = 1$)	$\tau^2+\sigma^2\left[1-\frac{(2\sqrt{\nu}d_{ij}\phi)^{\nu}}{2^{\nu-1}\Gamma(\nu)}K_{\nu}(2\sqrt{\nu}d_{ij}\phi)\right]$	262,743
Wave	$\tau^2+\sigma^2\left(1-\frac{\sin(\phi d_{ij})}{\phi d_{ij}}\right)$	262,087
Gaussian	$\tau^2+\sigma^2(1-\exp(-\phi^2 d_{ij}^2))$	262,039

Note: $N = 21{,}543$. Data from 2008 Cooperative Congressional Election Study. Notation from Banerjee, Carlin and Gelfand (2004, table 2.2). Bayesian information criterion (BIC) formula drawn from Gujarati and Porter (2009, 494). Computed with the geoR 1.7-4 library in R 3.0.0.

component of the model. For simplicity, label $\boldsymbol{\mu} = \mathbf{X}\beta+\mathbf{Zu}$ in (2) above and generalize the variance component according to:

$$\Sigma = \sigma^2\mathbf{H}(\phi)+\tau^2\mathbf{I}, \tag{3}$$

where σ^2 is a weight, called the *partial sill*, $\mathbf{H}(\phi)$ an $n\times n$ correlation matrix where the ijth value is defined by a parametric function, and τ^2 another weight, called the *nugget*, that multiplies an $n\times n$ identity matrix ($\tau^2+\sigma^2$ is called the *sill*). In this context, ϕ is called the *decay* parameter, and $1/\phi$ is called the *range* parameter. The ijth off-diagonal values of $\mathbf{H}$ actually could follow a variety of functional forms called *covariograms* (assumed spatial covariances at fixed distances), such as the wave, Gaussian, or Matérn functions, depending on what fits the data best (see Table 1 for a summary of some semivariograms). Often these parametric forms are given simple forms like $\exp(\phi d_{ij})$, an exponential correlation function, where the values d_{ij} are coordinate-based distance measures between case i and case j, weighted by autocorrelation parameter ϕ. These terms originate from Cressie (1991), and are further described in Banerjee, Carlin and Gelfand (2004).

The model is not yet complete, so modify it to specify a correlation at the group level:

$$\mathbf{Y} \mid \mu,\sigma^2,\tau^2,\phi \sim \mathcal{N}(\mathbf{X}\beta+\mathbf{Zu}+\mathbf{W},\tau^2\mathbf{I}), \tag{4}$$

thus adding a vector of *spatial random effects*:

$$\mathbf{W} = \{w(s_1), w(s_2), \ \ldots \ , w(s_n)\}, \tag{5}$$

such that each $Y(\mathbf{s}_i)$ is conditionally independent given the $w(\mathbf{s}_i)$ and $\mathbf{u}_i$, where $\mathbf{s}_i$ is a vector of location coordinates for observation i. The spatial aspect of the hierarchy is now given by:

$$\mathbf{W} \mid \sigma^2,\phi \sim \mathcal{N}(\mathbf{0},\sigma^2\mathbf{H}(\phi)), \tag{6}$$

where $H(\phi)_{ij} = \rho(\phi,\|\mathbf{s}_i-\mathbf{s}_j\|)$ and ρ a spatially appropriate correlation function. The spatial density blanket is now a function of $\mathbf{W}|\mathbf{Y}$, which naturally leads to credible intervals. Putting the hierarchical component together with the kriging component, along with a multivariate normal assumption, gives the full model. Hence, we now have a complete model of individual survey respondents that accounts for the respondents' locations in latitude and longitude. With this model in hand, we will use the technique of *point-to-block realignment* (spatial forecasts at new locations) to measure public sentiment for border-referenced constituencies (Banerjee, Carlin and Gelfand 2004, chapter 6). This model-based approach better accounts for dissimilarities across politically drawn lines and yields estimates with less measurement error as arbitrary boundaries like census tracts do not separate responses. Instead, the distance-based model recognizes shared effects by geography along with distinct effects by model hierarchy.

https://doi.org/10.1017/psrm.2015.5 Published online by Cambridge University Press

As a final point on background, Banerjee, Carlin and Gelfand (2004) give GeoBUGS code for specifying point-referenced models similar to what we have described. However, GeoBUGS is designed as an add-on package for WinBUGS, which only works on the Windows operating system. Further, GeoBUGS has proven to be computationally inefficient for these types of models and unable to handle large data sets effectively. Hence, we currently implement this structure using R.

Model for Unobserved Spatial Points

Our fundamental goal is to model the opinions of individuals, who are located in latitude and longitude, and use the information from this model to measure the ideology of a whole constituency. In this vein, Banerjee, Carlin and Gelfand (2004, 179–80) describe the method for using observations taken at various points in latitude and longitude and realigning them into measures of block-level averages. Their design assumes that the observed point-level data and the extrapolated block-level averages have a joint Gaussian distribution. If survey data include the latitude and longitude of survey respondents' residences (or, more commonly, information like ZIP code, for which coordinates can be assigned with reasonable accuracy), then point-to-block realignment offers an opportunity for effective measurement of precincts or congressional districts.

We want to create a posterior density blanket for some outcome of interest for the entire United States that can be sliced by any districting method where the partition is available. In the case of ideology, this would be an extension of Erikson, Wright and McIver (1993), where the boundaries are not required to be US states. This means that the final result will include estimates for spatial points where there is no observed data. Define now $\mathbf{s}$ as a set of n observed sites $\{\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_n\}$, where $Y(\mathbf{s})$ is an associated collection of outcomes $\mathbf{Y}(\mathbf{s}) = \{Y(\mathbf{s}_1), Y(\mathbf{s}_2), \ldots, Y(\mathbf{s}_n)\}$ and $\mathbf{X}^*(\mathbf{s}) = \{\mathbf{x}^*(\mathbf{s}_1), \mathbf{x}^*(\mathbf{s}_2), \ldots, \mathbf{x}^*(\mathbf{s}_n)\}$ is a collection of hierarchical covariates at those points. Note that the $\mathbf{X}^*(\mathbf{s})$ are conditional on the level of aggregation chosen (state, county, congressional district, etc.). Now alter the linear specification in (2) to accommodate this new setting:

$$\mathbf{Y}(\mathbf{s}) = \mu(\mathbf{s}) + \omega(\mathbf{s}) + \epsilon(\mathbf{s}), \tag{7}$$

where $\boldsymbol{\mu}(s) = \mathbf{X}(\mathbf{s})\boldsymbol{\beta} + \mathbf{Z}(\mathbf{s})\mathbf{u}$ is the mean structure based on a linear additive component, $\boldsymbol{\omega}(\mathbf{s})$ the realizations from a mean zero stationary (usually) Gaussian spatial process that captures spatial association (includes the partial sill σ^2 and the range ϕ), and $\boldsymbol{\epsilon}(\mathbf{s})$ a regular uncorrelated disturbance term (includes the nugget effect τ^2, its variance). In addition, assume that $\omega(s) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\sigma}^2\mathbf{H}(\phi))$ and $\epsilon(\mathbf{s}) \sim \mathcal{N}(\mathbf{0}, \tau^2\mathbf{I})$.

As noted, we usually specify the correlation function of $\boldsymbol{\omega}(\mathbf{s})$ to be a function of the separation between points to give a *stationary model*, so that:

$$\mu(\mathbf{s}) = E[Y(\mathbf{s})] \text{ and } f(\mathrm{Var}(\mathbf{s})) = f(\mathrm{Var}(\mathbf{s}+\mathbf{h})), \text{ for any } \mathbf{h} \in \Re^r,\ r \leq n, \tag{8}$$

where $\mathbf{h}$ is called the *separation vector*, and $f()$ an arbitrary user-defined function (which also just be the identity function). If this separation is purely based on distance, $\|\mathbf{s}_i - \mathbf{s}_j\|$, then we get *isotropy*, meaning that the *semivariogram* function

$$\gamma(\mathbf{h}) = \frac{1}{2}\mathrm{Var}(\mathbf{Y}(\mathbf{s}+\mathbf{h}) - \mathbf{Y}(\mathbf{s})), \tag{9}$$

depends on $\mathbf{h}$ only through its length $\|\mathbf{h}\|$. In this context, the nugget can be thought of as *micro-scale variability*: variability at distances smaller than the smallest inter-location distance in the data. This is important as we are now making predictions about points and distances between points that are not in the original data. Figure 2 gives an example of an exponential form.

https://doi.org/10.1017/psrm.2015.5 Published online by Cambridge University Press

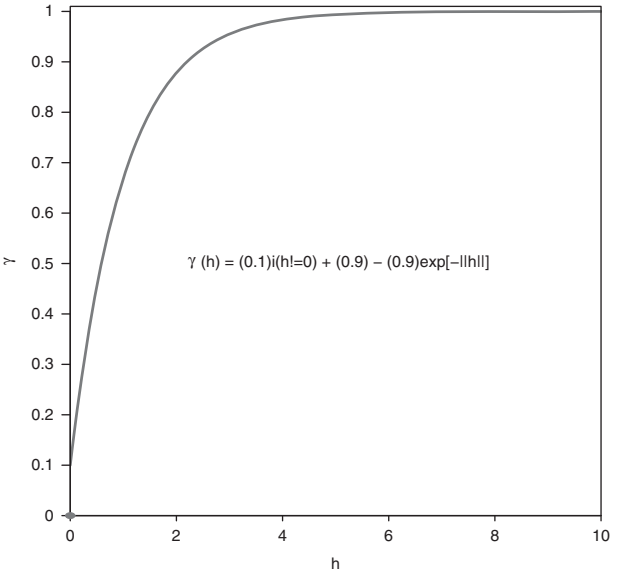


Fig. 2. Exponential semivariogram where $\tau^2 = 0.1$, $\sigma^2 = 0.9$, and $\phi = 1$

We clearly want the posterior distribution of the coefficient vector β for regression purposes, as well as the parameters σ^2, τ^2, and ϕ. However, we also want a smooth density blanket for $Y(\mathbf{s})$, which requires predictions at unobserved sites, generically labeled $\mathbf{s}_0$, giving $Y(\mathbf{s}_0)$ values at chosen locations. Kriging actually has two steps that are important here: the process of applying mathematical decision criteria, and then applying constrained optimization to get a density blanket. We consider this problem in the context of Gaussian processes, but many other distributional frameworks are available.

Optimization Criteria

The method of point-to-block realignment calls for kriging, or making forecasts of a spatially referenced variable (ideology, in our case) at new geographic locations. When choosing a kriging function with which to make forecasts, an obvious decision criterion is to find a function $f(\boldsymbol{Y})$ that minimizes the mean squared prediction error:

$$\begin{aligned} E[(Y(\mathbf{s}_0) - f(\mathbf{Y}))^2 \mid \mathbf{Y}] &= E[((Y(\mathbf{s}_0) - E[Y(\mathbf{s}_0) \mid \mathbf{Y}]) \\ &\quad + (E[Y(\mathbf{s}_0) \mid \mathbf{Y}] - f(\mathbf{Y})))^2 \mid \mathbf{Y}] \\ &= E[(Y(\mathbf{s}_0) - E[Y(\mathbf{s}_0) \mid \mathbf{Y}])^2 \mid \mathbf{Y}] \\ &\quad + E[(E[Y(\mathbf{s}_0) \mid \mathbf{Y}] - f(\mathbf{Y}))^2 \mid \mathbf{Y}]. \end{aligned}$$

However, $E\left[(E[Y(\mathbf{s}_0) \mid \mathbf{Y}] - f(\mathbf{Y}))^2 \mid \mathbf{Y}\right]$ has to be non-negative because of the square, so we know that

$$\begin{aligned} &E\left[(Y(\mathbf{s}_0) - E[Y(\mathbf{s}_0) \mid \mathbf{Y}])^2 \mid \mathbf{Y}\right] + E\left[(E[Y(\mathbf{s}_0) \mid \mathbf{Y}] - f(\mathbf{Y}))^2 \mid \mathbf{Y}\right] \\ &\geq E\left[(Y(\mathbf{s}_0) - E[Y(\mathbf{s}_0) \mid \mathbf{Y}])^2 \mid \mathbf{Y}\right]. \end{aligned}$$

https://doi.org/10.1017/psrm.2015.5 Published online by Cambridge University Press

The minimum (as in equality, not inequality above) happens only when $f(\mathbf{Y}) = E[Y(\mathbf{s}_0)|\mathbf{Y}]$, meaning $E[Y(\mathbf{s}_0)|\mathbf{Y}]$ is the function that minimizes mean squared prediction error. We know that $E[Y(\mathbf{s}_0)|\mathbf{Y}]$ is the *posterior mean* of $Y(\mathbf{s}_0)$ and therefore minimizes the posterior risk with a squared error loss function. Furthermore, an estimator that minimizes the Bayesian posterior risk is one that for *every* $Y \in \mathcal{Y}$ specifies a decision rule, $d(\mathbf{Y})$, that minimizes the posterior expected loss, $E_\pi[\mathbb{L}(A, d(\mathbf{Y}))]$, for action A. This is an overtly Bayesian statement of loss as we are averaging over uncertainty in the posterior distribution of $\boldsymbol{\beta}$, not over the *distribution* of data since it has already been observed. The subsequent decision rule which minimizes:

$$R_B(\boldsymbol{\beta}, d(\mathbf{Y})) = \int_B \int_{\mathcal{Y}} \mathbb{L}(A, d(\mathbf{Y}))dF_{\mathcal{Y}}(y \mid \boldsymbol{\beta})dF_\pi(\boldsymbol{\beta}), \tag{10}$$

for every $Y \in \mathcal{Y}$ is called *optimal*, and limiting the set to this choice is called a *Bayes rule*, denoted $\widehat{R_B}(\beta, d(\mathbf{Y}))$ here, and defined by the function minimizing value of β_A that satisfies:

$$\widehat{R_B}(\beta, d(\mathbf{Y})) = \inf_{\beta_A} \int_B R_F(A, d(\mathbf{Y}))dF_\pi(\beta), \tag{11}$$

(Gill 2014). This is then our criteria for finding $\boldsymbol{\beta}$ in (4). We describe a Markov chain Monte Carlo (MCMC) process for estimating the parameters in Appendix A, and this algorithm represents a large proportion of the contribution of this work.[2]

DATA APPLICATION: MEASURING PUBLIC IDEOLOGY IN 2008

Recall that our objective is a measure of public opinion (ideology and other public sentiment) for constituencies from both smoothed geographic terms as well as covariates. These are almost always areal or block units: states, congressional districts, state legislative districts, precincts, etc. Furthermore, survey data are typically sparse at the subnational level, and aggregated measures such as presidential vote share also pose problems (Kernell 2009). The best way to address these challenges is through an overtly hierarchical specification that references individual survey respondents, but adds covariate information at higher levels of aggregation (similar to past model-based methods, e.g., Pool, Abelson and Popkin 1965; Weber and Shaffer 1972; Weber et al. 1972; Jackson 1989; Park, Gelman and Bafumi 2006). Recognizing these differing levels in the model means that we can add extra information based on geographic location for *any* desired constituency.

Parameter Choices

As our model assumes the observed point-level data and the extrapolated block-level averages have a joint Gaussian distribution, we get:

$$f\left(\begin{pmatrix}\mathbf{Y_s}\\ \mathbf{Y_B}\end{pmatrix}\middle|\, \boldsymbol{\beta}, \sigma^2, \phi\right) = \mathcal{N}\left(\begin{pmatrix}\mu_{\mathbf{s}}(\boldsymbol{\beta})\\ \mu_{\mathbf{B}}(\boldsymbol{\beta})\end{pmatrix}, \sigma^2\begin{pmatrix}\mathbf{H_s}(\phi) & \mathbf{H_{s,B}}(\phi)\\ \mathbf{H}^T_{\mathbf{s,B}}(\phi) & \mathbf{H_B}(\phi)\end{pmatrix}\right),$$

where $\mathbf{Y_s}$ represents the vector of ideology among individual citizens, $\mathbf{Y_B}$ the vector of ideology in all block-referenced constituencies of interest, and $\mathbf{H}$ the correlation matrix of observations as before. Note that this presents the simplified case where there is no nugget effect (τ^2), but the

[2] A key advantage of this MCMC approach is it avoids estimation challenges outlined in Franzese and Hays (2007) by focussing on the post-converenge stationary distribution sampled by the process (Gill 2008).

https://doi.org/10.1017/psrm.2015.5 Published online by Cambridge University Press

result still holds if the variance–covariance terms do include a nugget, as seen in Appendix A. By standard normal theory (e.g., Ravishanker and Dey 2002), the conditional distribution of our extrapolated block averages is:

$$\mathbf{Y_B} \mid \mathbf{Y_s}, \boldsymbol{\beta}, \sigma^2, \phi \sim \mathcal{N}(\mu_{\mathbf{B}}(\boldsymbol{\beta}) + \mathbf{H}_{\mathbf{s},\mathbf{B}}^T(\phi)\mathbf{H}_{\mathbf{s}}^{-1}(\phi)(\mathbf{Y_s} - \mu_{\mathbf{s}}(\boldsymbol{\beta})),$$
$$\sigma^2[\mathbf{H_B}(\phi) - \mathbf{H}_{\mathbf{s},\mathbf{B}}^T(\phi)\mathbf{H}_{\mathbf{s}}^{-1}(\phi)\mathbf{H}_{\mathbf{s},\mathbf{B}}(\phi)]). \quad (12)$$

These quantities can be estimated with Monte Carlo integration:

$$(\hat{\mu}_{\mathbf{B}}(\beta))_k = L_k^{-1} \sum_{\ell} \mu(\mathbf{s}_{k\ell}; \boldsymbol{\beta})$$

$$(\widehat{\mathbf{H}}_{\mathbf{B}}(\phi))_{kk'} = L_k^{-1} L_{k'}^{-1} \sum_{\ell} \sum_{\ell'} \rho(\mathbf{s}_{k\ell} - \mathbf{s}_{k'\ell'}; \phi)$$

$$(\widehat{\mathbf{H}}_{\mathbf{s},\mathbf{B}}(\phi))_{ik} = L_k^{-1} \sum_{\ell} \rho(\mathbf{s}_i - \mathbf{s}_{k\ell}; \phi).$$

This will allow us to forecast the average ideology with:

$$\hat{\mu}_{\mathbf{B}}(\boldsymbol{\beta}) + \widehat{\mathbf{H}}_{\mathbf{s},\mathbf{B}}^T(\phi)\widehat{\mathbf{H}}_{\mathbf{s}}^{-1}(\phi)(\mathbf{Y_s} - \hat{\mu}_{\mathbf{s}}(\boldsymbol{\beta})). \quad (13)$$

We account for the spatial element by forecasting $\hat{Y}(\mathbf{s}_{k\ell}; \boldsymbol{\beta}, \sigma^2, \tau^2, \phi)$ and using this quantity in our Monte Carlo integration. With the nugget effect, from $\mathbf{Y(s)} = \boldsymbol{\mu}(\mathbf{s}) + \boldsymbol{\omega}(\mathbf{s}) + \boldsymbol{\epsilon}(\mathbf{s})$, we also get $\mathbf{Y_s} \sim \mathcal{N}(\boldsymbol{\mu}, \sum)$, where we still require $\sum = \sigma^2\mathbf{H}(\phi) + \tau^2\mathbf{I}$, $H(\phi)_{ij} = \rho(\phi, d_{ij})$ and $d_{ij} = \|\mathbf{s}_i - \mathbf{s}_j\|$. For this application, we selected the Gaussian semivariogram function for reasons shown below.

We want to forecast blocks using the conditional distribution in (12). However, current software is inadequate to jointly estimate the geospatial model and krige for more than a few kriged points, and there is no ability to obtain the conditional block averages. As a result, we developed our own sampler to more efficiently allow for kriging over a wide number of locations and more directly model the conditional mean of blocks (Appendix A).

Data for 2008

As our training data for this model we use the 2008 Cooperative Congressional Election Study (CCES) (Ansolabehere 2011). These data offer 21,543 observations spread across the American states and congressional districts, and the variables are described in Appendix B. Further, the data report ZIP codes, allowing us to place respondents in geographic space.[3] Figure 3 presents the geographic location of the respondents in the training data. The locations are presented on a scale of eastings and northings, which accounts for the spherical shape of the earth and places all observations on a common kilometer-referenced scale. By contrast, a degree of latitude need not constitute the same distance as a degree of longitude.[4]

The CCES presents a 100-point self-reported ideology scale, which serves as the outcome of interest.[5] It also provides covariates for age, education, race, sex, family income, religion,

[3] We do this by looking up the centroid of each ZIP code in the Census Bureau's gazetteer and then placing each respondent at a point jittered (adding a small amount of random noise) from the centroid but within the ZIP code's radius (United States Census Bureau 2002).

[4] The zero point for the eastings and northings scale is located at 37.5°N latitude and 96°W longitude, which is in southeast Kansas.

[5] The exact question wording from the CCES codebook: "One way that people talk about politics in the United States is in terms of left, right, and center, or liberal, conservative, and moderate. We would like to know how you view the parties and candidates using these terms. The scales below represent the ideological spectrum

https://doi.org/10.1017/psrm.2015.5 Published online by Cambridge University Press

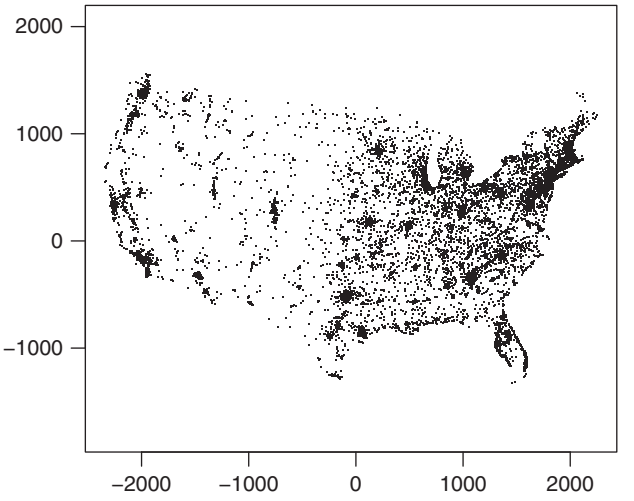


Fig. 3. Geographic locations in kilometer-referenced eastings and northings of survey respondents in the 2008 Cooperative Congressional Election Study

employment status, and homeowner status. We also include urban–rural status as a predictor by linking respondents' respective counties to the US Department of Agriculture's (USDA) scale for the relative ruralness of a county (USDA 2004). Lastly, with the eastings and northings generated by ZIP code, we include a polynomial function of location as a predictor.

When forecasting for states and congressional districts, we draw most covariates by sampling from the population distribution for the ZIP code in which the kriged point falls in, thereby drawing covariates consistent with the local population. We use the complete 2000 Census data, drawing from the joint distribution of age, race, and sex, and then drawing from the conditional distribution for all other predictors given these demographics. ZIP code-referenced Census data were archived with the National Historic Geographic Information System (Minnesota Population Center 2011).[6] For urban–ruralness, each kriged location in space is linked to the USDA county score. Moreover, as Census does not report religion, we independently draw from each county's distribution of religious adherents using the 2000 Religious Congregations and Membership Study (Jones et al. 2002; Finke and Scheitle 2005).[7] Thus, for a given kriged point, there is demographic information based on the ZIP code the point falls in, and we draw information on religiosity and ruralness based on the county in which the point falls.

Model Specification and Estimation

In addition to our predictors of age, education, race, sex, income, religion, employment status, homeowner status, and urban–rural status, our model of mean ideology also includes a spatial

(F'note continued)

from very liberal (0) to very conservative (100). The most centrist American is exactly at the middle (50). Where would you place yourself? If you are not sure, or don't know, please check 'Not Sure.'"

[6] As an alternative analysis, we also made our forecasts using complete observations from the 2000 Census 1 percent unweighted sample by state, thereby accounting for the joint distribution of the predictors at the state level. These data were obtained from IPUMS (Ruggles et al. 2010). Our results were similar to those reported in their work.

[7] These data were downloaded from the Association of Religion Data Archives (http://www.TheARDA.com) and were collected by the Association of Statisticians of American Religious Bodies.

https://doi.org/10.1017/psrm.2015.5 Published online by Cambridge University Press

trend term. The spatial trend surface is a linear combination of polynomials of the spatial coordinates. Following the prescription of Cressie (1993, 155–6), the polynomial terms are defined by the formula:

$$\sum_{0 \leq k+\ell \leq r} \sum \alpha_{k\ell} s_{i1}^{k} s_{i2}^{\ell}, \quad \mathbf{s}_i = (s_{i1}, s_{i2})', \tag{14}$$

where $\mathbf{s}_i$ is the vector location of observation i, s_{i1} the easting for observation i, s_{i2} the northing for observation i, α_{kl} the coefficient for the term where the easting is raised to the exponent k and the northing is raised to the exponent l, and r the order of the polynomial trend term. In short, for an rth-order trend surface, every product term for which the exponents on eastings and northings sum to any quantity $\leq r$ is included. After comparing the fit of models including a constant trend ($r = 0$) up through a fourth-order trend ($r = 4$) using Schwarz's Bayesian information criterion (BIC), we found that the second-order and third-order models fit the best. Owing to multicollinearity issues, simplicity in reporting, and a slightly better penalized fit, we present the results of the second-order model ($r = 2$).[8]

With this linear structure in place for the model of ideology's mean, the full Bayesian model can be specified as follows:

$$\begin{aligned}
\mathbf{Y_s} &\sim \mathcal{N}(\mathbf{X_s}\boldsymbol{\beta}, \Sigma) \\
\Sigma &= \sigma^2 \mathbf{H}(\phi) + \tau^2 \mathbf{I} \\
H(\phi)_{ij} &= \rho(\phi; d_{ij}) \\
\rho &= \exp(-\phi^2 d_{ij}^2) \text{ (Gaussian correlation function)} \\
\pi(\boldsymbol{\beta}) &\sim \text{flat} \\
\pi(\tau^2/\sigma^2) &\sim \text{Unif}(6, 8) \\
\pi(\sigma^2) &\sim 1/\sigma^2 \\
\pi(1/\phi) &\sim \text{Unif}(0, 9000).
\end{aligned} \tag{15}$$

In this setup, rather than modeling the nugget effect (τ^2) directly, we model the relative nugget (τ^2/σ^2) or ratio of the nugget to the directly modeled partial sill. As a reasonably tight prior is required for numerical stability (Hobert and Casella 1996), we centered a uniform distribution around an ~7:1 empirical ratio of these two quantities. For the partial sill (σ^2), we assume a reciprocal prior, and for the range ($1/\phi$), we assume a uniform prior up to twice the maximum distance in the data. Priors on the regression coefficients are flat. The best-fitting covariance function is the Gaussian function, which will be described more in the next section, and all other model attributes are defined as before.

To estimate this model, we adapt a five-step algorithm developed by Diggle and Ribeiro (2002, 141).[9] First, we draw several values from a discrete version of the uniform priors

[8] The procedure for choosing a spatial trend term is similar to using polynomials to model the baseline hazard rate in an event history analysis (Box-Steffensmeier and Jones 2004, 75–6). Specifically, a good fit of a smoothed intercept is the goal. In this case, for the second-order model, the BIC was 65.37, whereas the next best BIC was 65.40 for the third-order model. These scores are remarkably similar, but the second-order model scores slightly better, and the third-order model was inestimable owing to multicollinearity in the Bayesian context. Hence, we chose to estimate the second-order model.

[9] See also Diggle and Ribeiro (2007, chapter 7).

https://doi.org/10.1017/psrm.2015.5 Published online by Cambridge University Press

for $\frac{\tau^2}{\sigma^2}$ and $\frac{1}{\phi}$. Second, we estimate the conditional posterior distribution, $p\left(\frac{\tau^2}{\sigma^2}, \frac{1}{\phi} \mid \mathbf{Y}\right)$ by placing our draws from the discrete prior into the following formula:

$$p\left(\frac{\tau^2}{\sigma^2}, \frac{1}{\phi} \mid \mathbf{Y}\right) \propto \pi\left(\frac{\tau^2}{\sigma^2}\right)\pi\left(\frac{1}{\phi}\right) \mid V_{\tilde{\beta}} \mid^{\frac{1}{2}} \left|\mathbf{H}(\phi) + \left(\frac{\tau^2}{\sigma^2}\right)\mathbf{I}\right|^{-\frac{1}{2}} (\hat{\sigma}^2)^{-\frac{n}{2}}, \tag{16}$$

where $V_{\tilde{\beta}}$ is the correlation matrix of the regression coefficients estimated with generalized least squares (GLS) using the current draw of $1/\phi$, n the sample size, and $\hat{\sigma}^2$ an estimate of the partial sill based on residuals drawn from the GLS coefficient estimates.[10] All other terms are defined as before. Third, we draw a single set of sample posterior values for $\frac{\tau^2}{\sigma^2}$ and $\frac{1}{\phi}$ from (16). Fourth, we attach the set of sampled values to $p\left(\boldsymbol{\beta}, \sigma^2 \mid \frac{\tau^2}{\sigma^2}, \frac{1}{\phi}, \mathbf{Y}\right)$ and compute the corresponding conditional posterior distributions as:

$$\sigma^2 \mid \mathbf{Y}, \frac{\tau^{2^*}}{\sigma^2}, \frac{1^*}{\phi} \sim \chi^2_{ScI}(n, \hat{\sigma}^2)$$

$$\boldsymbol{\beta} \mid \mathbf{Y}, \sigma^2, \frac{\tau^{2^*}}{\sigma^2}, \frac{1^*}{\phi} \sim \mathcal{N}(\tilde{\boldsymbol{\beta}}, \sigma^2 V_{\tilde{\beta}}). \tag{17}$$

Here, the term χ^2Scl refers to the Gaussian-Scaled-Inverse-χ^2 distribution. The terms in these equations are again drawn from the GLS estimates from the initial draw of ϕ. After taking a draw from the scaled inverse χ^2 distribution for the partial sill σ^2, this term is linked with the draws from the relative nugget and range terms when drawing the regression coefficients from a normal distribution. By repeating the third and fourth steps to generate a sufficiently large sample from each of the conditional posteriors, we build a sufficient Monte Carlo sample to reflect the joint posterior for the full parameter set $\left(\frac{\tau^2}{\sigma^2}, \frac{1}{\phi}, \sigma^2, \boldsymbol{\beta}\right)$.

RESULTS

Appendix C reports the results of a simple linear model of self-reported ideology on a 0 (liberal) to 100 (conservative) scale, estimated with ordinary least squares (OLS). This model previews our structural perspective of modeling ideology as a function of demographics recorded in Census data. Following Park, Gelman and Bafumi (2004), we include the predictors of age, education, race, and sex. We also add income, religion, ruralism, homeownership, and employment status as demographic predictors. (Hence, all of these are individual-level variables except for ruralism, which is a county-level variable.) Per the discussion of Equation (14), we also include a second-order polynomial trend of ideology based on the survey respondent's location in eastings and northings. This aspect allows the ideology mean to vary by region in a globally defined way. Our subsequent model will make the assumption that, once this trend has been removed, additional localized spatial clustering can be captured through a model of the error variance.

Figure 4 presents the empirical semivariogram of our data. The horizontal axis represents the distance separating survey respondents in kilometers, and the vertical axis reports the semivariance among observations spread that far apart. In this figure, the distance between two observations is the Euclidean distance based on the two-dimensional coordinates of eastings and northings: each point on the figure represents the estimated semivariance among a bin of observations separated by approximately the distance on the horizontal axis. The semivariance can be interpreted as either half the variance of the *differences* in the outcome over a certain

[10] Specifically, $\tilde{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{H}(\phi)^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{H}(\phi)^{-1}\mathbf{Y}$. Hence, $V_{\tilde{\beta}} = (\mathbf{X}'\mathbf{H}(\phi)^{-1}\mathbf{X})^{-1}$. In addition, $\hat{\sigma}^2 = \frac{1}{n}(\mathbf{Y}-\mathbf{X}\boldsymbol{\beta})'\mathbf{H}(\phi)^{-1}(\mathbf{Y}-\mathbf{X}\boldsymbol{\beta})$.

https://doi.org/10.1017/psrm.2015.5 Published online by Cambridge University Press

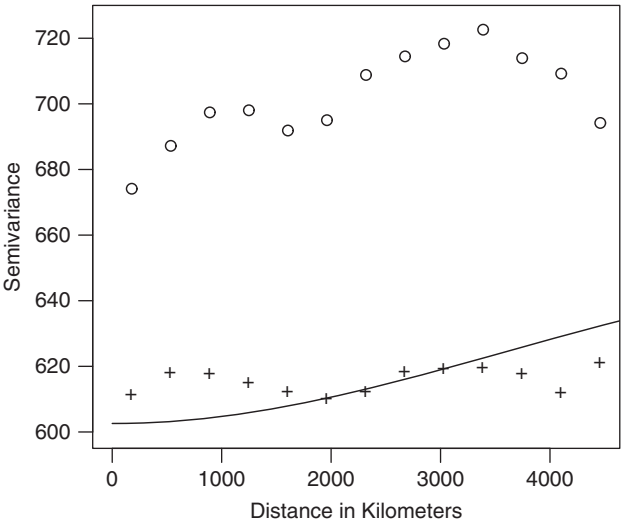


Fig. 4. Empirical semivariogram of respondent ideology in 2008 Cooperative Congressional Election Study
Note: Circles show the empirical semivariogram for unmodeled ideology, plus signs show the empirical semivariogram for residuals of ideology from a linear model, and the line shows the fitted Gaussian semivariogram from a jointly estimated Bayesian model.

separation of length ||**h**|| (e.g., $\frac{1}{2}Var(Y(s+h)-Y(s))$), or the full variance of the *level* of the outcome given this distance (e.g., Var(**Y**| ||**h**||)) (Matheron 1963; Cressie and Hawkins 1980). The open points represent the semivariance for the raw, unmodeled observations of ideology. The plus signs represent the semivariance for the residuals from a linear model estimated with OLS. Naturally, the variance of the residuals is consistently smaller owing to some variance being explained by the model.

As is expected of spatially referenced data, more proximate observations have smaller variance, indicating that they are more highly correlated. As distance increases, the semivariance increases as the observations become more independent of each other. As a comparison, the semivariance of unmodeled observations with no separation is ~675, meaning that the variance in ideology scores among respondents with no spatial separation is 675. By contrast, among observations separated by ~3300 km, the semivariance is over 720: this means that the variance in ideology scores among respondents separated by 3300 km is 720, and the variance of the differences, $\mathbf{Y(s+h)-Y(s)}$, when ||**h**|| = 3300 is 1440. In using this graph to choose a parametric semivariogram in specifying the complete spatial model, we compare the patterns in the raw semivariogram to the empirical signature produced by several common parametric forms.[11] We observed that a Gaussian semivariogram function best fit the pattern of the raw data, and wave function looked more appropriate for the residuals.

Table 1 describes the four commonly used semivariogram specifications that we considered for our analysis. The wave, Matérn, and Gaussian functions were the most plausible because these three allow the decay in spatial correlation to accelerate at intermediate distances. (This is exemplified in Figure 4 by the fact that the variance stays fairly low up to around 2000 km, then

[11] This process of determining the proper form for spatial correlation is similar to the process of specifying an ARMA model for time-dependent correlation using the signature produced in autocorrelation and partial autocorrelation functions (Enders 2009, 68).

https://doi.org/10.1017/psrm.2015.5 Published online by Cambridge University Press

rapidly rises.) As a point of reference, we also estimate the functionally simple and widely used exponential semivariogram function. As the table shows the Gaussian model had the lowest BIC value, indicating the best fit. The exponential, by contrast, fit far worse than the other three candidates. As the simplest and best-fitting model, we proceed with the Gaussian semivariogram function.

With this sense of the structure of spatial correlation, we turn to a model that estimates the mean and error structure simultaneously. Current software is limited in its ability to process massive amounts of spatial data, so we introduce the technique of *Bootstrapped Random Spatial Sampling* (BRSS) to estimate the full model. We proceed by drawing 100 subsamples with replacement from the full data. Each sample consists of 5 percent of the original data. We then run the Monte Carlo sampler described in the previous section on each subset, save 1000 iterations, and pool the results across all 100 subsamples. Appendix C reports the results of this estimation with posterior means, standard errors, and credible intervals from the Monte Carlo integration process.[12] Many of the results are substantively similar to the OLS findings from before. However, the estimates of the coefficients have become more diffuse. In addition, the estimates of the variance terms differ somewhat, with a smaller nugget effect and range term.[13] The posterior mean for the nugget effect (τ^2) indicates that the non-spatial unexplained variance is around 605, whereas the posterior mean for the partial sill (σ^2) indicates that the spatial unexplained variance is around 87. With ~12.6 percent of the unexplained variance accounted for by space, spatial smoothing appears to add a substantial degree of explanatory power.

Again, Figure 4 presents the original unconditional semivariogram and the semivariogram of OLS residuals. The line in this figure presents the Gaussian semivariance function using the posterior means of the BRSS. The Bayesian model simultaneously estimates the regression coefficients and spatial covariance parameters, so the reported line ought to better reflect the spatial correlation. Relative to the empirical semivariogram of OLS residuals reported in the figure, we see a stronger correlation at lower distances and a weaker one at farther distances, which is theoretically sound.

Forecasts for the Lower 48 States and DC

We proceed to estimate the ideology of the lower 48 states and Washington, DC in four steps. First, we draw points from across the United States, link each point with its corresponding state, and pair the point with a random draw of demographic data from the 2000 Census for the relevant ZIP code. When simulating citizens, we sample first by ZIP code in proportion to the population of the ZIP code, then randomly draw a location around the ZIP code's centroid. In this manner, the geographic distribution of our simulated points reflects the geographic distribution of population in the United States. During this stage, we create 241,530 simulated citizens who are assigned a specific location in latitude and longitude as well as all covariate values. Second, for every iteration of our Monte Carlo sample, we draw 5 percent (or 12,076) of these simulated voters with replacement. Third, we use the parameter values at the given

[12] All BRSS results and the resulting forecasts were computed using the Georgia Advanced Computing Resource Center's Linux cluster (zcluster).

[13] Though the average of the range term is just outside the maximum distance in the data at 4955 km, the semivariogram switches from convex to concave within the domain of the data. This means that the spatial correlation starts to approach its lower bound at $\frac{1}{\sqrt{2\phi}} = 3504$ km. The location of the convex to concave switch can be deduced by taking the second derivative with respect to d of the semivariogram function shown in Table 1, where d is distance, the second derivative switches from positive to negative at $d = \frac{1}{\sqrt{2\phi}}$.

https://doi.org/10.1017/psrm.2015.5 Published online by Cambridge University Press

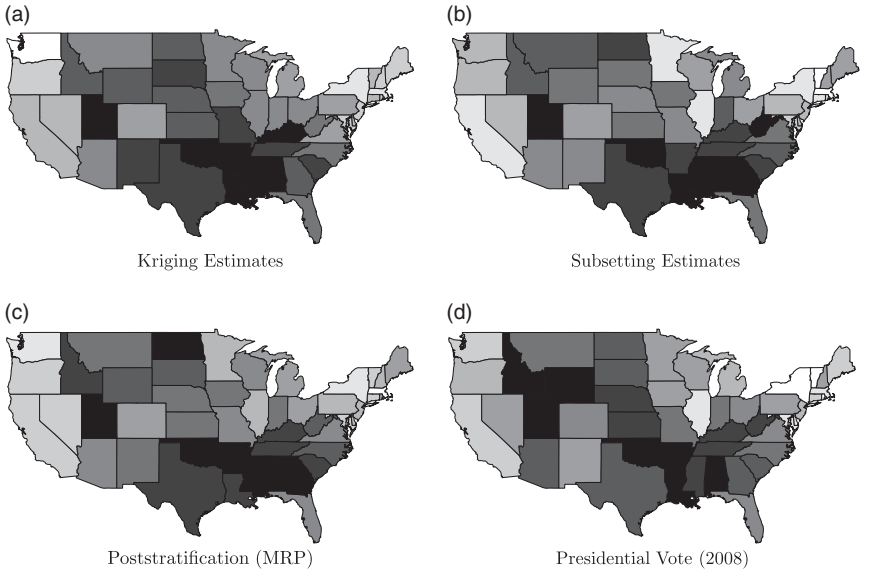


Fig. 5. Three estimates of ideology and presidential vote by state for the lower 48 states and DC
Note: Darker values mean more conservative (a–c) and a below-average vote for Obama (d). Lighter values mean more liberal and an above-average vote for Obama.

iteration to forecast ideology for the 12,076 simulated citizens with the 5 percent of the data used to fit the particular iteration. Fourth, after kriging ideology for every iteration of the sample, we average the ideology of all of the kriged points within a state.

We present our estimates in Figure 5. Figure 5(a) shows a map of our estimates, with darker values indicating that citizens in the state are more conservative than in the average state and lighter values indicating that they are more liberal. As a contrast to our estimate, Figure 5(b) shows the findings we obtain if we simply subset the original CCES data by state and take the average ideology of all respondents.[14] Similarly, Figure 5(c) shows the estimates that can be generated using multi-level regression with post-stratification (sometimes called MRP for "multi-level regression and post-stratification").[15] Finally, Figure 5(d) shows Obama's share of the two-party vote in 2008. States shaded darker were states where McCain performed better than average, and lighter-shaded states were where Obama turned in his best performances. The pattern in all four maps of Figure 5 is roughly similar. The biggest difference is that the estimates from kriging put an emphasis on global trends so states are more likely to look similar to each other than they do with the other measures. Despite the empirical regularity that partisan polarization is aligning on regional lines, it is possible that including a second-order spatial trend in the model of mean ideology is forcing one more degree of neighbor similarity than is warranted.

As has been mentioned, an added benefit of measuring ideology with kriging, or point-to-block realignment, is that measures of the variance in the concept are straightforward. Figure 6 illustrates

[14] The reported results do not post-stratify with survey weights, as the unweighted measures fit better. Weighted measures, however, are available on our Dataverse page.

[15] Functionally, the MRP model is the same as the model presented in Appendix C, except there are no spatial trend or spatial variance terms. Instead, a state-level random intercept is included. Forecasts are based on covariate draws from the 2000 Census data.

https://doi.org/10.1017/psrm.2015.5 Published online by Cambridge University Press

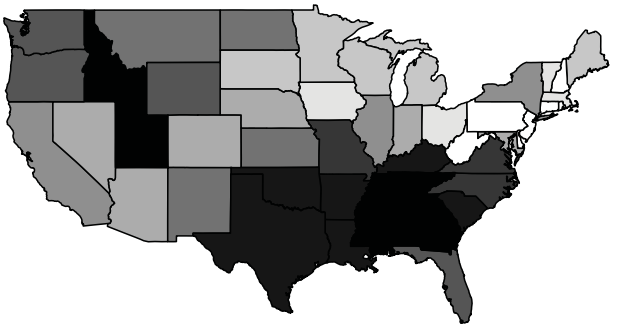


Fig. 6. Variance of ideology as forecasted by kriging
Note: Darker values mean greater variance, whereas lighter values mean lesser variance.

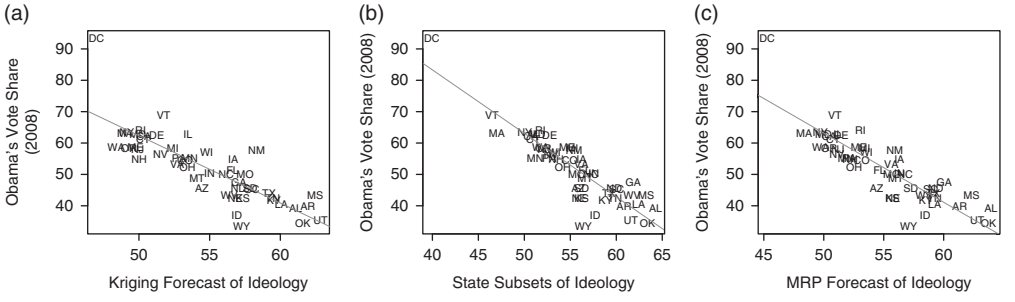


Fig. 7. Scatterplot of 2008 presidential vote by state against three measures of ideology: (a) kriging, (b) subsetting, (c) and post-stratification

this by showing a map of the estimated variance in ideology. Darker states have a higher variance in ideology, whereas lighter states have a lower variance in ideology. As might be expected, the most variable states have the most demographically diverse populations, whereas more homogeneous states have a tighter variance.

The comparison and contrast between kriging, simple subsets, and post-stratification is further illustrated in scatterplots using the three measures. Figure 7 shows a scatterplot of Obama's share of the two-party vote against the kriged measure of ideology, the subsetted measure, and the post-stratification forecast. Each panel also includes a simple regression line fit to the illustrated data. As would be expected, on all measures the more conservative a state's electorate is ideologically, the lower Obama's share of the two-party vote was.[16] The correlation coefficients were all large and negative, though kriging actually lags slightly in explained variance ($r = -0.8912$ for the subsetting measure, $r = -0.8668$ for post-stratification, and $r = -0.8281$ for the kriging measure). This is seen visually in Figure 7(b), which shows the steepest regression line for the subsetting measure, with less dispersion around the line. Hence, when the luxury of a survey sample of tens of thousands is available and reasonable sample subsets can be obtained for each state, the subsetting estimates of ideology appear to work best. We turn now, however, to a case where large subsamples are not feasible.

[16] An alternative means of evaluating these measures would be to correlate them with measures of state legislative ideology (Shor and McCarty 2011).

https://doi.org/10.1017/psrm.2015.5 Published online by Cambridge University Press

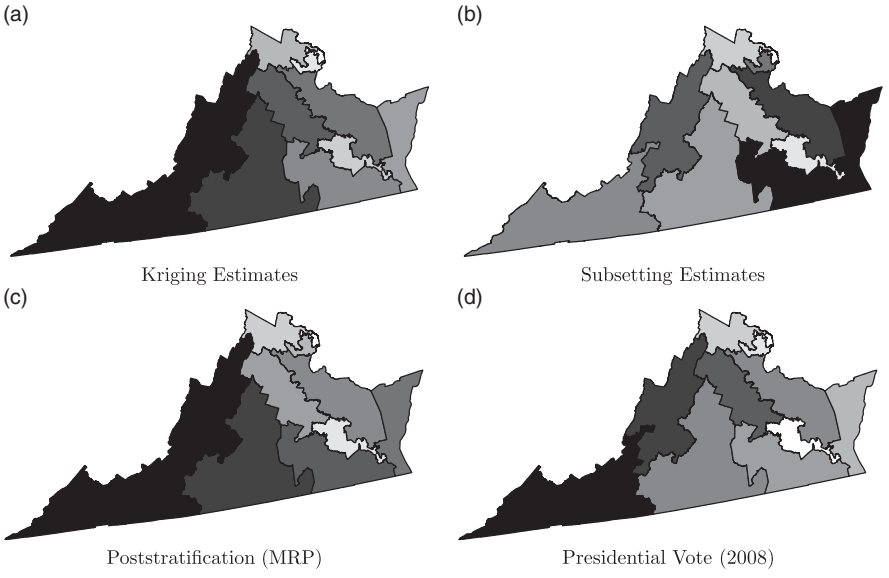


Fig. 8. Three estimates of ideology and presidential vote by district for Virginia's congressional districts in 2008

Note: Darker values mean more conservative (a–c) and a below-average vote for Obama (d). Lighter values mean more liberal and an above-average vote for Obama.

Forecasting Congressional District Ideology: The Case of Virginia

To illustrate the flexibility of our technique, we now turn to a geographically narrower constituency, congressional districts. It is rare to find survey data from the same polling house that allows the user to create subsets of respondents by congressional district in order to obtain comparable estimates of electoral ideology or public opinion of any sort. In fact, the CCES is unusually good in this regard as part of its design was to allow users to be able to estimate quantities for states and congressional districts. Even in this case, though, subsamples for districts can become quite small: in the case of Virginia, the fifth district had 33 respondents in the CCES, the 11th district had 65 respondents, and all other districts fell somewhere in between.

We use the case of Virginia as an example of how kriging can be used to forecast constituency ideology in congressional districts as it is an interesting state in this regard: a strong geographic contrast between the DC suburbs and the rest of state as well as a history of racial divide. We use the same process as was described for forecasting state ideology. This time, we use a subsample of 48,720 kriged points specific to Virginia's 11 congressional districts, which is a subset of the broader sample of 241,530 kriged points nationwide. Figure 8 presents our kriged estimates of ideology, the estimates from simple averaging of CCES respondents in each district, post-stratification forecasts of ideology, and Obama's share of the two-party vote in 2008. Just as before, darker-shaded districts indicate that the district has above-average conservatism or a below-average Obama share of the two-party vote, whereas lighter-shaded districts indicate above-average liberalism or an above-average Obama share of the two-party vote. Again, the patterns are somewhat similar among the methods. This time it looks as if the kriged estimates correspond particularly well to the presidential vote by district.

Figure 9 offers further verification that our kriged estimates serve as a good predictor of the presidential vote. This figure shows a scatterplot of the kriged estimate of ideology, the district

https://doi.org/10.1017/psrm.2015.5 Published online by Cambridge University Press

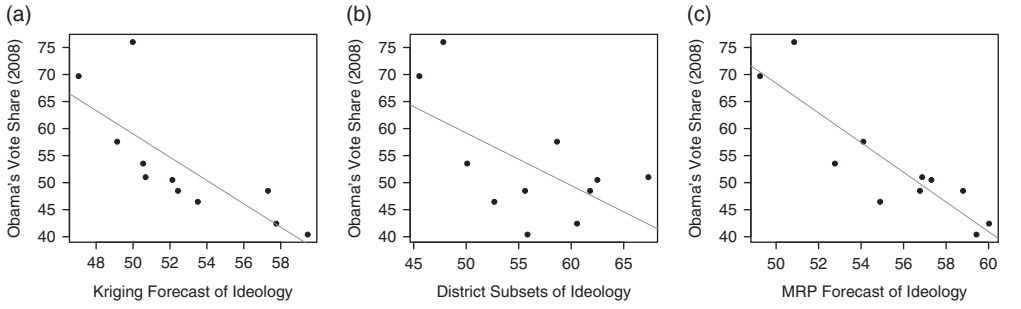


Fig. 9. Scatterplot of 2008 presidential vote in Virginia's congressional districts against three measures of ideology: (a) kriging, (b) subsetting, (c) and post-stratification

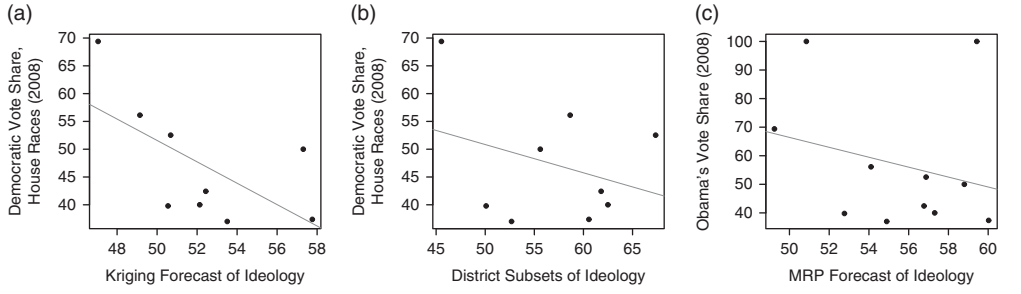


Fig. 10. Scatterplot of 2008 congressional vote in Virginia's congressional districts against three measures of ideology: (a) kriging, (b) subsetting, (c) and post-stratification

subsample estimates, and post-stratification estimates, each against Obama's share of the two-party vote in the district. Although there are only 11 observations to consider, the regression lines are somewhat steeper, and the residuals are somewhat smaller, for the model-based methods of kriging and post-stratification. The latter actually serves as the best predictor of presidential vote ($r = -0.8852$), but kriging also performs well ($r = -0.7750$). Subsetting, with a small sample size, now correlates much more weakly with presidential vote ($r = -0.6001$).

As a final look at this measure, we examine the Democratic candidate's share of the two-party vote in the 2008 election for the House of Representatives seat in each of these districts. Only nine of these seats were contested, so we only examine these cases. Figure 10 again shows scatterplots of the Democratic candidate's share of the two-party vote against each of the three measures of constituency ideology. As Figure 10(a) shows, the slope of the regression line is steepest for the kriging estimates, and the residuals are smaller than for the other two measures. Although there are only nine observations here, the kriging measure fits somewhat better than the post-stratification measure ($r = -0.6270$ for kriging and $r = -0.6018$ for post-stratification), whereas the subsetting measure performs noticeably worse than the others ($r = -0.3213$). Overall, then, the kriging measure performs strongly in this application of measuring ideology in congressional districts.

CONCLUSION

Our contribution here is twofold. First, we have made a theoretical and empirical argument that Bayesian hierarchical spatial modeling is an effective means of measuring public opinion in

https://doi.org/10.1017/psrm.2015.5 Published online by Cambridge University Press

defined geographic constituencies. This methodology improves forecasting by minimizing the completely stochastic variance in models of citizen ideology, and using the results of such a model to reasonably project citizen preferences into their constituencies of residence. A central advantage of this model-based means of measurement is that it can be applied to any locale in time and space: given a snapshot of geographically referenced data at a given point in time, our model can be used to generate estimates of ideology for any border-referenced constituencies of interest. The advantage here extends further in that a forecast can be generated for a constituency, even if there are no training observations from the constituency (e.g., no observations in smaller states or in congressional districts). This contrasts from multi-level regression with post-stratification, where some training data in each constituency is essential to obtaining a prediction of the constituency's random effect.

Second, we have developed software to implement the Bayesian kriging approach. Our objective is to provide easy-to-use software for researchers to tessellate national ideology at any level of geographic division (state, congressional district, county, etc.). Software solutions available currently impose severe limitations for such techniques: the requirement to consider all dyads of observed data to estimate spatial correlation and the computational complexity of numerically integrating this model create substantial memory and calculation demands. However, the Monte Carlo integration routines we describe in this paper, when implemented in R and GeoBUGS, are limited in their ability to fit spatial models or krige over new locations only by the user's computational resources, mostly RAM as R swaps to disk frequently during large calculations.

Despite software challenges, we were able to develop the method of BRSS that can be estimated over a computing cluster in the span of a few days using a modest multi-blade cluster. These results yielded similar findings in state-level and congressional district-level ideology as simple subsets of the CCES and multi-level regression with post-stratification, but with better variance properties. As additional studies are incorporated into the software solution and more comprehensive full-information MCMC methods are developed, this advantage will increase. Furthermore, our solution provides better coverage in geographic areas where the original data are sparse, as the kriging estimation for these areas borrows strength from neighboring regions in exactly the same way that small groups borrow strength from large groups at the aggregation level of a standard hierarchical model.

REFERENCES

Ansolabehere, Stephen. 2011. 'CCES, Common Content, 2008, Version 4'. Available at http://hdl.handle.net/1902.1/14003, accessed April 18, 2013.

Ansolabehere, Stephen D., James M. Snyder, and Charles Stewart. 2001. 'Candidate Positioning in U.S. House Elections'. *American Journal of Political Science* 45:136–59.

Banerjee, Sudipto, Bradley P. Carlin, and Alan E. Gelfand. 2004. *Hierarchical Modeling and Analysis for Spatial Data*. New York, NY: Chapman & Hall.

Berry, William D., Evan J. Ringquist, Richard C. Fording, and Russell L. Hanson. 1998. 'Measuring Citizen and Government Ideology in the American States, 1960–93'. *American Journal of Political Science* 42:327–48.

Beyle, Thad, Richard G. Niemi, and Lee Sigelman. 2002. 'Gubernatorial, Senatorial, and State-Level Presidential Job Approval Ratings: The U.S. Officials Job Approval Ratings (JAR) Collection'. *State Politics and Policy Quarterly* 2:215–29.

Box-Steffensmeier, Janet M., and Bradford S. Jones. 2004. *Event History Modeling: A Guide for Social Scientists*. New York, NY: Cambridge University Press.

https://doi.org/10.1017/psrm.2015.5 Published online by Cambridge University Press

Brace, Paul, Kevin Arceneaux, Martin Johnson, and Stacy G. Ulbig. 2004. 'Does State Political Ideology Change Over Time?' *Political Research Quarterly* 57:529–40.

Brace, Paul, Kellie Sims-Butler, Kevin Arceneaux, and Martin Johnson. 2002. 'Public Opinion in the American States: New Perspectives Using National Survey Data'. *American Journal of Political Science* 46:173–89.

Cohen, Jeffery E. 2006. *Public Opinion in State Politics*. Stanford, CA: Stanford University Press.

Cressie, Noel A. C. 1991. *Statistics for Spatial Data*. New York, NY: Wiley.

——. 1993. *Statistics for Spatial Data*, rev. ed. New York, NY: Wiley.

Cressie, Noel A. C., and Christopher K. Wikle. 2011. *Statistics for Spatio-Temporal Data*, 1st ed. New York, NY: Wiley.

Cressie, Noel A. C., and Douglas M. Hawkins. 1980. 'Robust Estimation of the Variogram: I'. *Mathematical Geology* 12(2):115–25.

DeLeon, Richard E., and Katherine C. Naff. 2004. 'Identity Politics and Local Political Culture: Some Comparative Results from the Social Capital Benchmark Survey'. *Urban Affairs Review* 39(6): 689–719.

Diggle, Peter J., and Paulo J. Ribeiro Jr. 2002. 'Bayesian Inference in Gaussian Model-Based Geostatistics'. *Geographical & Environmental Modeling* 6(2):129–46.

——. 2007. *Model-Based Geostatistics*. New York, NY: Springer.

Djupe, Paul A., and Anand E. Sokhey. 2011. 'Interpersonal Networks and Democratic Politics'. *Political Science and Politics* 44(1):55–59.

Elazar, Daniel J. 1966. *American Federalism: A View from the States*. New York, NY: Thomas Y. Crowell.

Enders, Walter. 2009. *Applied Econometric Time Series*, 3rd ed. New York, NY: Wiley.

Erikson, Robert S., and Gerald C. Wright. 1980. 'Policy Representation of Constituency Interests'. *Political Behavior* 2:91–106.

Erikson, Robert S., Gerald C. Wright, and John P. McIver. 1993. *Statehouse Democracy: Public Opinion and Policy in the American States*. New York, NY: Cambridge University Press.

——. 2006. 'Public Opinion in the States: A Quarter Century of Change and Stability'. In Jeffrey E. Cohen (ed.), *Public Opinion in State Politics*, 229–53. Stanford, CA: Stanford University Press.

Finke, Roger, and Christopher P. Scheitle. 2005. 'Accounting for the Uncounted: Computing Correctives for the 2000RCMS Data'. *Review of Religious Research* 47(1):5–22.

Fischer, David Hackett. 1989. *Albion's Seed: Four British Folkways in America*. New York, NY: Oxford University Press.

Franzese, Robert J. Jr., and Jude Hays. 2007. 'Spatial Econometric Models of Cross-Sectional Interdependence in Political Science Panel and Time-Series-Cross-Section Data'. *Political Analysis* 15 (2):140–64.

Garreau, Joel. 1981. *The Nine Nations of North America*. Boston, MA: Houghton Mifflin.

Gastil, Raymond D. 1975. *Cultural Regions of the United States*. Seattle, WA: University of Washington Press.

Gelman, Andrew, and Jennifer Hill. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. New York, NY: Cambridge University Press.

Gelman, Andrew, and Thomas C. Little. 1997. 'Poststratification into Many Categories Using Hierarchical Logistic Regression'. *Survey Methodology* 23:127–35.

Gill, Jeff. 2008. 'Is Partial-Dimension Convergence a Problem for Inferences from MCMC Algorithms?' *Political Analysis* 16(2):153–78.

——. 2014. *Bayesian Methods for the Social and Behavioral Sciences*, 3rd ed. New York, NY: Chapman & Hall/CRC.

Gill, Jeff, and Andrew J. Womack. 2012. 'The Multilevel Model Framework'. In Brian Marx, Marc Scott, and Jeff Simonoff (eds), *The Sage Handbook of Multilevel Modeling*, 3–20. Thousand Oaks, CA: Sage.

Gimpel, James G., and Jason E. Schuknecht. 2003. *Patchwork Nation: Sectionalism and Political Change in American Politics*. Ann Arbor, MI: University of Michigan Press.

Gray, Virginia. 1976. 'Models of Comparative State Politics: A Comparison of Cross-Sectional and Time Series Analysis'. *American Journal of Political Science* 20:235–56.

https://doi.org/10.1017/psrm.2015.5 Published online by Cambridge University Press

Gujarati, Damodar N., and Dawn C. Porter. 2009. *Basic Econometrics*, 5th ed. New York, NY: McGraw-Hill/Irwin.

Hobert, James P., and George Casella. 1996. 'The Effect of Improper Priors on Gibbs Sampling in Hierarchical Linear Mixed Models'. *Journal of the American Statistical Association* 91:1461–

Huckfeldt, Robert, and John T. Sprague. 1995. *Citizens, Politics, and Social Communication: Influence in an Election Campaign*. New York, NY: Cambridge University Press.

Jackson, John E. 1989. 'An Errors in Variables Approach to Estimating Models with Small Area Data'. *Political Analysis* 1:157–80.

——. 2008. 'Endogeneity and Structural Equation Estimation in Political Science'. In Janet M. Box-Steffensmeier, Henry E. Brady and David Collier (eds), *Oxford Handbook of Political Methodology*, 404–31. New York, NY: Oxford UniversityPress.

Jackson, John E., and David C. King. 1989. 'Public Goods, Private Interests, and Representation'. *American Political Science Review* 83(4):1143–

Jacoby, William G., and Saundra K. Schneider. 2001. 'Variability in State Policy Priorities: An Empirical Analysis'. *Journal of Politics* 63:544–68.

Jones, Dale E., Sherry Doty, Clifford Grammich, James E. Horsch, Richard Houseal, Mac Lynn, John P. Marcum, Kenneth M. Sanchagrin, and Richard H. Taylor. 2002. *Religious Congregations and Membership in the United States 2000: An Enumeration by Region, State and County Based on Data Reported for 149 Religious Bodies*. Nashville, TN: Glenmary Research Center.

Kellsall, J., and P. J. Diggle. 1995. 'Non-Parametric Estimation of Spatial Variation in Risk'. *Political Analysis* 14:2335–

Kernell, Georgia 2009. 'Giving Order to Districts: Estimating Voter Distributions with National Election Returns'. *Political Analysis* 17:215–35.

Lawson, A. B., and F. Williams. 1993. 'Application of Extraction Mapping in Environmental Epidemiology'. *Statistics in Medicine* 12:1249–

Lax, Jeffrey R., and Justin H. Phillips. 2009. 'How Should We Estimate Opinion in the States?' *American Journal of Political Science* 53:107–21.

Lieberman, Robert C., and Greg M. Shaw. 2000. 'Looking Inward, Looking Outward: The Politics of State Welfare Innovation Under Devolution'. *Political Research Quarterly* 53:215–40.

Lieske, Joel. 1993. 'Regional Subcultures of the United States'. *Journal of Politics* 55(4):888–913.

Matheron, G. 1963. 'Principles of Geostatistics'. *Economic Geology* 58:1246–266.

Minnesota Population Center. 2011. *National Historic Geographic Information System: Version 2.0 [Machine-Readable Database]*. Minneapolis, MN: University of Minnesota.

Norrander, Barbara. 2001. 'Measuring State Public Opinion with the Senate National Election Study'. *State Politics and Policy Quarterly* 1:113–27.

Pacheco, Julianna 2011. 'Using National Surveys to Measure Dynamic U.S. State Public Opinion: A Guideline for Scholars and an Application'. *State Politics and Policy Quarterly* 11:415–39.

Park, David K., Andrew Gelman, and Joseph Bafumi. 2004. 'Bayesian Multilevel Estimation with Poststratification: State-Level Estimates from National Polls'. *Political Analysis* 12:375–85.

——. 2006. 'State-Level Opinions from National Surveys: Poststratification Using Multilevel Logistic Regression'. In Jeffrey E. Cohen (ed.), *Public Opinion in State Politics*, 209–28. Stanford, CA: Stanford University Press.

Pool, Ithiel de Sola, Robert P. Abelson, and Samuel L. Popkin. 1965. *Candidates, Issues, and Strategies*. Cambridge, MA: MIT Press.

Putnam, Robert D. 1966. 'Political Attitudes and the Local Community'. *American Political Science Review* 60(3):640–54.

——. 1993. *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton, NJ: Princeton University Press.

Ravishanker, Nalini, and Dipak K. Dey. 2002. *A First Course in Linear Model Theory*. Boca Raton, FL: Chapman & Hall/CRC.

Ruggles, Steven, J. Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek. 2010. *Integrated Public Use Microdata Series: Version 5.0 [Machine-Readable Database]*. Minneapolis, MN: University of Minnesota.

https://doi.org/10.1017/psrm.2015.5 Published online by Cambridge University Press

Shor, Boris, and Nolan McCarty. 2011. 'The Ideological Mapping of American Legislatures'. *American Political Science Review* 105(3):530–51.

Sinclair, Betsy. 2012. *The Social Citizen: Peer Networks and Political Behavior*. Chicago, IL: University of Chicago Press.

Tam Cho, Wendy K., and James G. Gimpel. 2007. 'Prospecting for (Campaign) Gold'. *American Journal of Political Science* 51(2):255–68.

Tausanovitch, Chris, and Christopher Warshaw. 2013. 'Measuring Constituent Policy Preferences in Congress, State Legislatures, and Cities'. *Journal of Politics* 75(2):330–42.

United States Census Bureau. 2002. 'Census 2000 Gazetteer Files: Zip Code Tabulation Areas'. Available at http://www.census.gov/geo/maps-data/data/gazetteer2000.html, accessed April 20, 2013.

United States Department of Agriculture (USDA). 2004. '2003 Rural–Urban Continuum Codes'. Available at http://www.ers.usda.gov/data-products/rural-urban-continuum-codes.aspx, accessed July 24, 2013.

Weber, Ronald E., Anne H. Hopkins, Michael L. Mezey, and Frank Munger. 1972. 'Computer Simulation of State Electorates'. *Public Opinion Quarterly* 36:549–65.

Weber, Ronald E., and William R. Shaffer. 1972. 'Public Opinion and American State Policy-Making'. *Midwest Journal of Political Science* 16:683–99.

https://doi.org/10.1017/psrm.2015.5 Published online by Cambridge University Press