

Political Science and History: Enhancing the Methodological Repertoire

Gregory J. Wawro*

Associate Professor

Department of Political Science

gjw10@columbia.edu

Ira Katznelson

Ruggles Professor of Political Science and History

Departments of Political Science and History

iik1@columbia.edu

Columbia University

420 W. 118th St.

New York, NY 10027

phone: 212-854-8540

fax: 212-854-5670

*Corresponding author. The authors gratefully acknowledge the research assistance of Quinn Mulroy, Grant Porter, Josh Savitch, Tiffany Washburn, and Steven White.

Political Science and History: Enhancing the Methodological Repertoire

Abstract

Seeking to advance historical studies of political institutions and behavior, we argue for an expansion of the standard methodological toolkit with a set of innovative approaches that privilege parameter heterogeneity to capture nuances missed by more commonly used models. So doing, we address critiques by prominent historians and historically-oriented political scientists who have underscored the shortcomings of mainstream quantitative approaches for studying the past. They are concerned that the statistical models ordinarily employed by political scientists are inadequate for addressing temporality, periodicity, specificity, and context—issues that are central to good historical analysis. The innovations in method that we advocate are particularly well-suited for incorporating these issues in empirical models, which we demonstrate with two replications of extant research. The replications focus on locating structural breaks associated with reform of legislative rules, and on the temporal evolution in congressional roll-call behavior regarding labor policy during the New Deal and Fair Deal.

1 Introduction

How should the craft, logics, and landscape of history be made a constitutive aspect of political analysis? This question and the complex of issues that inform it have become pressing as more and more scholarship in political science has been taking a historical tack, even in parts of the discipline that long had been oriented to the present or the very recent past. Historians and historical social scientists often blanch at what they believe to be far too simple—and simplifying—ways of working on the past by mainstream colleagues (Gaddis 2002; Sewell 2005; Tractenberg 2006; Tilly and Goodin 2006). Demonstrating how some recent advances in quantitative modeling can engage with, and help resolve, their often on-target criticisms, this article contributes to the systematic engagement of political science with historical materials, and rejects any choice between historical depth and methodological sophistication.

A growing number of ambitious accounts trace how processes, including the temporal horizons of actors, unfold over time. We can see this development in studies of public opinion (Page and Shapiro 1992; Berinsky 2009) and the character and control of regulatory agencies over the long term in the United States (Moe 1987; Carpenter 2010). Though not wholly new as a research program (Brady 1973; Cooper and Brady 1981), this trend is especially pronounced in studies of Congress (Binder 1997; Schickler 2000; Wawro and Schickler 2006; Jenkins, Peck, and Weaver 2010). Such scholarship also has become more common among quantitative and formal students of international relations (Mansfield and Snyder 2005) and comparative politics (Kalyvas 1996, 2006). Attention to history, the discipline has learned, offers a rich reservoir of information that can be utilized by deductive and quantitative, as well as a range of qualitative, methods, sometimes serving as material with which to test models and hypotheses that first were broached in analyses of more current situations (Bates, Greif, Levi, Rosenthal, and Weingast 1998; King, Keohane, and Verba 1994).

But therein often lies a problem. History frequently is treated simply as data with the implicit assumption of an equivalence among facts and particulars found in various periods and contexts. History, the critics rightly insist, is more than a mere laboratory or depository of additional cases for model building and testing.

They have two primary concerns. Substantively, they think this emergent work misses the

chance to enlarge scholarship’s agenda, to ask the most important, difficult, and challenging historical questions. Methodologically, they find this scholarship to be naive about the challenges that come with taking history seriously. Without attention to analytical issues posed by temporality, periodicity, specificity, and context, the utilization of historical data, they argue, paradoxically raises barriers to good historical scholarship, and thus, in the absence of sufficient methodological self-consciousness, may actually widen a long-standing breach between history and political science (Katznelson and Lapinski 2006).

Wishing to promote this engagement, we seek to build on how some historical institutionalists have been ruminating about how to place “politics in time,” while mounting their own criticism of the mainstream discipline for its neglect or mechanical deployment of the past, and for an apparent lack of interest in the most significant large questions about the character of the polity—issues that only can be addressed with a historical sensibility (Pierson 2004; Orren and Skowronek 2004; Pierson and Skocpol 2002).

Convinced that there is a tremendous opportunity waiting to be seized, we argue for a set of innovations in methodology that can complement what less formal and more qualitative researchers can accomplish, while taking seriously the central themes and epistemological concerns articulated in their critiques and their scholarship. While others have convincingly pointed to flaws in quantitative approaches to history, the literature lacks constructive suggestions for how to augment existing methods in order to overcome the shortcomings they have identified. Our aim is to demonstrate how methods that were in the main not developed for historical analysis can offer exactly what critics of such modeling claim is lacking.

The article proceeds as follows. We provide an overview of various criticisms that concern how political science approaches history, proceed to a discussion that contrasts more standard quantitative approaches with promising alternatives, and offer two sets of replications to illustrate their plausibility and advantages. We conclude by suggesting ways to move this research program forward.

2 Skepticism and Possibility

Many historians pay little or no attention to political science, thinking its practitioners to be primitive or uncurious about the past. On this view, political scientists spend too

little time with primary sources in archives or with secondary sources written by historians. But these are relatively superficial matters; after all, there is no reason political scientists cannot exercise more due diligence with historical evidence or broaden their libraries (Kreuzer 2010). But some historians who take the social sciences seriously, and who wish, like Richard Hofstadter writing a half-century ago, to develop history “informed by the insights of the social sciences,” whose “primary purpose will be analytical” (Hofstadter 1956, 363), have posed more fundamental challenges. Such historians know that the two crafts, irrespective of an overlap, have distinctive purposes and qualities, and they understand that political scientists, within the division of scholarly labor, are less focused on period-specific narratives of “what happened” than attuned to theoretical and empirical models fashioned to discern causes and mechanisms.

Their concerns run both ways. An example is William Sewell’s *Logics of History* (2005). Just as he worries about social scientists who apply models over broad swaths of history in ways that risk ignoring changes to context, temporality, and historical specificity, he laments how historians often do not attend to theory and methodology, substituting a “narrative overconfidence” in which the story can mask the absence of a systematic approach to causation or interpretation (Sewell 2005, 11).

Sewell’s main contribution is to ask us to think more richly about types of temporality. He distinguishes three kinds. These he calls teleological, in which outcomes are the result of large-scale processes, beyond events and often beyond agency, like modernization, state-building or capitalist development; experimental, in which time is seen as equivalent and the comparative method is used to tease out causal parallels across time; and eventful, sensitive to the context specific intertwining of structure and agency (Sewell 2005, 83). Interestingly defining “events” not as anything that happens, but “as that relatively rare subclass of happenings that significantly transform structures,” he stresses “a causal dependence of later occurrences on prior occurrences and assumes that social causality is temporally heterogeneous, not temporally uniform” (Sewell 2005, 100–101). All three approaches to time, he says, can be productive, but only if what he calls the eventful approach is integrated with the others. This kind of history must be particularly attentive to how categories of understanding and analysis emerge and change at different moments, as meanings are not static

(73ff). From this vantage, negotiation of choices are informed not only by particular contexts and situations but also by particular tools of meaning and understanding that actors have that are embedded in particular historical situations.

Located at the border between respect and dismissal, John Lewis Gaddis's *Landscapes of History* (2002) offers an even more challenging assessment by a historian seeking to engage with the social sciences. Though overstated, his complaints deserve to be taken seriously as condensed and concentrated versions of reservations held by historians. Gaddis sets out from the premise that apprehending history as a whole is impossible. As they seek to connect causes to consequences, he counsels, historians thus depend on approaches and methods that can move back and forth between more literal and more abstract representations of the past, and between generalization and particularization (Gaddis 2002, 12–15). This is exactly the challenge that confronts historically-oriented political science, a challenge he depicts in a metaphor of the tension between maps that depict reality and mapping as an act of choosing how to simplify in order to show what is real. Such acts perform the work of modeling, and they range from more to less simplified and from more to less logical as distinct from more to less empirical. History, from this perspective, depends not only on more or better facts, but on the quality of its methods and the character of its analytical reason.

Gaddis believes that many social scientists (as opposed to non-laboratory natural scientists) fall short in how they conduct such exercises, and it is to these issues that we should attend, despite his disconcerting tendency to caricature too broadly. His concerns fall into three main categories.

The first relates to the role of particularity and context when dealing with variables that possess a logic of tendency that can be generalized across time and space. To illustrate his point, he uses the example of mountains. There is not an endless variety of types or shapes; all mountains are constrained by a logic of geometry. Notwithstanding, each actual mountain is particular and specific (Gaddis 2002, 83). History, too, has a limited range and trans-historical features. But, he insists, particularity matters. “Causes always have contexts, and to know the former we must understand the latter” (Gaddis 2002, 97).

From this point of view, context is not a kitchen-sink category, but a feature of reality that can be tied closely to analyses of causality. Defined as “the dependency of sufficient

causes upon necessary causes,” Gaddis treats context as those features of social reality in specific locations and moments that make a given cause a candidate for sufficient status (Gaddis 2002, 97). Taken seriously, this implies that any specific factor or variable only can gain such potency in context, not outside it. This is a powerful cautionary lesson. It does not do away with more general claims about the logic of things, like the geometry of mountains, but it insists that this logic alters as parameters change.

Second, Gaddis is unhappy with what he represents, in exaggerated form, as the propensity of social scientists to treat dependent variables in terms of rich variation awaiting explanation by a discovery of the key independent variable. He wishes to invert this balance of parsimony and complexity, strongly preferring work that identified a single, highly-targeted object of analysis coupled to an acknowledgment that causation is often complex and interdependent.

Third, he observes that temporality matters causally. He insists not only that there is a constellation of causes, but that we have to be sensitive to their heterogeneity in time, noting the “distinctions that have to be made ... between the immediate, the intermediate, and distant” (Gaddis 2002, 95). This, in fact, is a point of emphasis rather similar to Sewell’s characterization of the causal effects of time and what Pierson tries to accomplish in *Politics in Time* in focusing on path dependency, heterogeneous causality, and contingency in an effort described as that of joining “structure, conjuncture, and events in a common causal universe” (Pierson 2004, 102, 109).

Taken together, the three concerns expressed by Gaddis invite political scientists to think more systematically about how, in Sewell’s terms, to bring “a serious infusion of historical habits of mind” into theory and empirical inquiries. It invites us, further, to read and extract lessons sympathetically from how historians, and historically-oriented social scientists contextualize causal accounts, and what they do regarding particularity, periodization and temporality, and how they treat the fatefulness and importance of sequence, the layering of processes, variations to rates of change, and the configurational contingency of events. So it is to relevant methodological questions that we now turn to suggest how we might develop a response to the call for more “sophistication about temporality” (Sewell 2005, 6).

3 Methodological issues

A major problem that historians and scholars in the field of American Political Development (APD) have with causal empirical modeling is the manner in which a reduction in complexity inherent in such endeavors is carried out. Of course, some historians reject causal modeling out of hand, claiming that their mandate is simply to explain what, not why, things happen. But others, as we have seen, are much more sympathetic and attuned to the powers of social science, yet remain uncomfortable with how modeling of complex reality, especially quantitative modeling, proceeds.

This, we believe to be wholly explicable. After all, the workhorse quantitative method in political science is regression analysis. By its very nature, the regression model is an attempt to represent complex parts of the world in a highly parsimonious way by taking a number of dimensions (i.e., a set of explanatory variables) and reducing them to one (i.e., the dependent variable). Some historians reject the notion that the world can be expressed in terms of such simple relationships. This criticism applies not just to dimensionality reduction, but also to the direction of causality. It is difficult for historians to accept that there are variables that are truly exogenous to others.

Although there are well-known methods for addressing issues of endogeneity and social scientists have recently become more focused on unearthing more persuasive and powerful instrumental variables for identifying causal effects (Angrist and Pischke 2009), in practice the assumptions necessary for the methods to perform well are often not met and the promise falls well short of what many hope for (Deaton 2009). These methods can involve complex systems of equations where certain variables are allowed to explain as well as be explained. While such systems can be easy to write down, data limitations may simply prevent empirical identification of the direction of causal arrows. We think that more progress can be made, however, by augmenting standard models with techniques that have only recently come to the attention of political scientists. These techniques work within the framework of the regression model in ways that satisfy a desire for parsimony while incorporating features historians see as crucial to representing historical events and development.

A key feature of an alternative approach is to posit more general models than are commonly employed and to rely more heavily on the data to tell us about the model's structure.

Historians often feel uncomfortable with the restrictive confines of a regression model. At some level, these kinds of restrictions are unavoidable. But we can incorporate more flexibility by having researchers impose less *a priori* structure on the phenomena under investigation.

The central features of history that simple regression models fail to capture are temporality, periodicity, specificity, and context. Temporality refers to the importance of the sequence of events, periodicity to the clustering of events along the dimension of time, specificity to the uniqueness of events in historical time, and context to larger environments within which events take place. These concepts overlap to a degree: context may vary according to period, for example. Thus, understanding events may require grouping them into particular phases or epochs despite perpetual uncertainty about the nature of the periodization and the multiplicity of plausible ways to divide historical time.

The main reason that standard regression models are poorly equipped to address these concerns is that they typically treat parameters as non-varying over long stretches of history. It is not uncommon for political scientists to estimate a regression model over decades, even centuries, where the relationship among variables is treated as constant over the entire period.¹ Coefficients on explanatory variables—which capture how several dimensions reduce to one—do not vary and thus assume away a significant amount of complexity in the development and evolution of historical processes. While such simple models are consistent with the desire to derive and test general theories, they raise an alarm for scholars who doubt the tremendous—perhaps excessive—determinism that such models imply. While statistically significant relationships may give us confidence that we have discovered a grand and fundamental pattern of behavior or an institutional truth, such models tend to impose far too much structure on data, concealing important nuances and thereby providing unsatisfying and perhaps even incorrect accounts. We may draw the wrong inferences about historical processes if the relationship between variables is nonexistent in certain periods, but is exceptionally strong for others.

Parameter variation offers a potentially powerful solution. It permits the effects of ex-

¹For example, see Binder 1997; Brady, Buckley, and Rivers 1999; Brunell and Grofman 1998; Schickler 2000; Wawro and Schickler 2006. We are, of course, in no way claiming that this is “bad” work. All we are saying is that historians may find the quantitative approaches adopted in these studies as lacking sensitivity to some of their concerns.

planatory variables to change and evolve along dimensions that we deem important. For historical analysis, time is generally the most important dimension. Sometimes we are interested in the values of parameters before and after some watershed event—a so-called structural break. Other times we might be interested in how parameters might change gradually but decidedly over time. In either case, the estimation approach that we use must allow parameters on variables of interest to change over time, even as the values of those variables may be changing over time. Note that we are not arguing that political scientists should forsake parsimony in their desire to incorporate complexity. We are sympathetic to arguments that the fewer variables in a model the better and that more explanatory variables do not necessarily lead to better models or more thorough analysis. Yet few variables with varying parameters over time can accomplish a surprisingly nuanced account that is adequately sensitive to historical complexity.

Suppose we want to analyze annual data over a broad swath of history. Time series analysis of the kind typically conducted in political science would likely estimate one coefficient value for a given explanatory variable for the entire series. Flexibility of the kind historians believe to be essential can be introduced by letting that coefficient vary over the series. But letting parameters be different for each observational unit in the series (i.e., each year), is not desirable either. Such a move makes it impossible for a univariate time series to identify annual coefficients, since it would yield as much, or more, parameter variation as variation in the data. Nor would this degree of parameter variation always be preferable, since it may very well leave us with results that are difficult if not impossible to interpret in a way that is useful for hypothesis testing or as a means to make any sort of general claim.

Yet it is possible to locate an attractive middle ground, especially if we have data variation along another dimension beside time. A key concern is that as the number of parameters that we are estimating increases, it becomes harder to estimate them with satisfactory precision, since less data is being brought to bear on the estimation of each parameter.² We might mistakenly infer variation in the relationships among variables simply because we are unable to obtain very precise estimates of parameters. Estimation strategies that are data-dependent in terms of the structure of models need to strike a balance between the amount

²See Bartels 1996 for a nice discussion of the trade-offs with pooling data and parameter precision.

of information that is used to identify varying parameters from the data and the amount of structure imposed by the researcher. What would be helpful practically and theoretically is to allow coefficients in a regression model to vary while linking them together in ways that can capture temporality, periodicity, specificity, and context.

We suggest that historically-oriented political scientists could employ two sets of methods to accomplish this: structured additive regression models and flexible methods for modeling structural change. Structured additive regression (STAR) models extend generalized additive models by incorporating flexible nonparametric functions of covariates that can account for nonlinear effects and build in complexities in relationships among variables that are not possible in standard regression models (Fahrmeir and Tutz 2001; Hastie and Tibshirani 1990). Several classes of models familiar to political scientists, including generalized additive mixed models, variable coefficient models, and multilevel/hierarchical models, are special cases of STAR models. STAR models can capture parameter variation as well as unobserved heterogeneity that is likely to exist in the data while performing parameter smoothing to reduce the estimation instability that can result when we increase the parameter-to-data ratio.³ Smoothing can be done in a way that is particularly sensitive to arguably the most important dimension in historical studies—time. But smoothing can also be used to provide flexibility over other dimensions as well, such as region and policy area.

Structural change models—also known as change point models—allow for different parameter values for distinct parts of the data. Structural breaks are a key concern for APD scholarship; (usually exogenous) significant events occur and relationships among variables are subsequently markedly different. Studies that allow for a single structural break in the data are quite common in political science. The standard approach is to include a dummy variable indicating the time period when the break is believed to occur and interact that dummy with relevant variables of interest. This approach has serious drawbacks, however,

³Isaac and Griffin (1989) claim that those who incorporate parameter variation in historical quantitative analysis “will have to develop a tolerance for ... large sampling variances of the estimators,” and consequently “reduced statistical significance” of estimates. Bayesian approaches help to address this concern by reducing variability in the parameter estimates. Smoothing is also often referred to as “shrinkage”, where we are “shrinking” free parameter estimates back to some common mean. Partial pooling in multilevel models is another type of smoothing approach (Gelman and Hill 2007).

since it does not allow much flexibility in determining either where the break actually occurred or whether more than one break exists in the data. We seek a more flexible approach in both determining the existence and location of structural breaks as well as estimating coefficients in different periods demarcated by the breaks.

Bayesian approaches to estimation of STAR models and structural changes offer the kinds of flexibility that we seek. One of the key advantages to Bayesian approaches is that we can use priors that incorporate assumptions that are particularly useful for historical analysis. These assumptions pertain to how we think relationships among variables might play out over the time dimension, and help us manage the trade-off when increasing the parameter-to-data ratio.⁴

For example, suppose we were examining the effects of different factors on the roll call votes of members of Congress on civil rights measures over the course of the 19th and 20th centuries. Parameters on member-specific characteristics could be permitted to vary annually or by congress if we believe that different factors may have different effects over time. Or we may have specific periodization schemes where parameters would be constrained to be similar across specific years. While we could impose these constraints explicitly, a STAR model can be employed with Bayesian estimation approaches that would enable us to have the data determine to a large extent which periodization scheme is most appropriate.

Particular kinds of priors offer solutions to the problem of pooling observations over long stretches of history. Complete pooling of the data is the kind of transgression that historians refer to when they complain that political scientists ignore the “texture and complexity” of history (Silbey 2000, 326). Allowing the effects to vary across the time dimension of the data builds in some complexity that complete pooling would ignore. Yet, of course, we cannot

⁴Priors also offer a systematic way to incorporate the rich information provided by historians’ scholarship. However, the quantification of qualitative accounts to form priors can be difficult. Our interest lies in the structure of priors, especially with respect to time, more than in the specification of particular parameter values in priors. While we generally adopt a utilitarian position when it comes to methods, the philosophical foundation as well as instrumental application of Bayesian approaches may appeal more to historians than frequentist statistics. While the frequentist notion of taking repeated samples often does not make sense for studies of contemporary political behavior (e.g., when our data is the universe of cases), it makes even less sense when studying behavior in a specific time period with a unique historical context (cf. Western and Jackman 1994).

let parameter effects vary completely or we risk explaining nothing. We also want to avoid overstating the variance of parameter estimates.

Bayesian estimation of semi-parametric models can help us strike a balance between incorporating more of the complexity that historians like to see and imposing the kind of structure that is necessary to model behavior quantitatively with the goal of revealing underlying patterns. For example, we could conduct the roll call analysis mentioned above using a hierarchical version of a STAR model where congresses are grouped according to the party systems within which they occurred. It is quite likely that roll call voting behavior will vary under different party systems given how parties line up along issue cleavages. Allowing parameter effects to vary across congresses in this way can help account for periodicity and variation in context. If we believe that particular variables are related to period-specific or contextual variation, we can model parameters as functions of these variables.

The downside to doing this, of course, is that we could get the periodization scheme or contextualization wrong—whether by specifying it in terms of measurable variables or not—which could result in misspecification bias. Undoubtedly, there will be disagreement over any explicitly imposed structure on parameter variation. In cases where such disagreement is particularly contentious, we could choose to let periodization or contextualization schemes be substantially data-dependent. One way to do this is to employ directed and undirected autoregressive priors, which could be used to smooth parameters such that the effects of variables will be more similar for adjacent time periods (Breslow and Clayton 1993).⁵

For example, in a roll call analysis, immediately adjacent congresses could be designated as similar without specifying exactly the nature of the similarity apart from proximity. Let C_t denote a congress at time t . C_{t+1} and C_{t-1} would be designated most similar to C_t . The similarity would decay as we move away from t , so that we would posit that $C_{t\pm 2}$ would be less like C_t than $C_{t\pm 1}$, and so on. In this way, parameters for a given congress would be estimated by “borrowing strength” from proximate congresses without imposing an explicit periodization scheme. Borrowing strength in this manner enables parameter variation without giving up too much in terms of the precision of estimates.⁶ As the parameters

⁵For other examples of parameter smoothing priors in political science, see Bartels 1996.

⁶While correlation among observations is often viewed as a nuisance in quantitative analysis, here we seek to turn it into an advantage.

are smoothed across time periods, periodization/contextualization schemes would emerge dynamically based on what is occurring in the data. It is important to keep in mind that these are priors. They do not force parameters to be similar; they are merely “suggestions” that the estimation technique uses to find parameter estimates. By defining hyperpriors on variance parameters, the amount of smoothness can be estimated simultaneously with the regression coefficients to see what patterns emerge in terms of which observational units are more similar than others. Splines along the time dimension could be employed here to smooth over temporally varying parameters and capture complex non-linearities in the effects of explanatory variables (Brezger and Lang 2006).

The modeling of proximity need not be constrained to the time dimension. Priors that link geographic or policy areas could be adopted, drawing on methods from spatial statistics to introduce this kind of complexity. In this sense, we approach the modeling historical processes by thinking of them more abstractly as maps (cf. Gaddis 2002). This permits a multidimensionality that should be particularly appealing for those who wish to see more complexity in the models deployed to study history. To continue the roll call voting example, in addition to incorporating priors that account for proximity over the time dimension, we could expand these priors to model proximity in terms of type of vote—for example, on amendments, on final passage, on substantive issues, on procedural matters. The effects of variables could be allowed to vary both across time and type of vote, leading to a proliferation of parameters that would be made manageable by the priors. A good deal of progress has been made recently in the theory and implementation of intrinsic autoregressive (IAR) priors, which are particularly useful for dealing with this type of parameter variation. IAR priors are a type of Markov random field (MRF) prior, which set up a general functional relationship for the parameters for different observational units that captures the similarity or proximity of the units along different dimensions.⁷

While periodization schemes may emerge from this kind of approach, periodicity may be investigated more directly by focusing on models for structural breaks. Structural breaks are a key concern for APD scholarship; (usually exogenous) significant events occur and rela-

⁷Initial work on MRF priors was undertaken by Besag 1974 (see also Besag 1975; Besag and Kooperberg 1995; Besag, Green, Higdon, and Mengersen 1995). For an accessible discussion on IAR priors, see Girosi and King 2008.

tionships among variables are subsequently different. Determining the existence of structural breaks and locating when they occur can be tricky with conventional methods. A standard approach to testing whether an event is historically important or not is simply to include a dummy variable distinguishing observations before and after the occurrence of the event. But simply inserting the dummy in a particular time period can be problematic. Suppose the event occurs at time t , but we do not observe the change until time $t + 2$. Does this mean that the event was not important? Or if another event occurs at time $t + 2$, it is entirely possible that a dummy variable indicating observations before and after time t will have a (misleading) statistically significant coefficient.

The textbook method for testing for a break in a time series is the Chow test. This test assumes that it is valid to break a series into two parts—before and after some posited significant event. Yet it could be the case that the series should be broken into more than two parts, and that the most important breaks occur at different points in time from those previously assumed. If there are more structural breaks in the data than specified, we could reach incorrect inferences. Standard tests of structural breaks require a specification of where they occur. In some cases it is better to be agnostic, and let the data tell us how many breaks there are and where they occur.⁸

Several approaches are available that do not have the restrictions of commonly employed methods. Western and Kleykamp (2004) develop a Bayesian changepoint model for univariate time series that does not specify the location of the change *a priori*, and instead treats its location as a parameter to be estimated. The inferences about regression coefficients produced from the model account for uncertainty about the changepoint—another departure from conventional models. Chib (1998) develops a multiple changepoint model that posits regimes as hidden states from which observations are drawn. An unobserved discrete state variable, indicating which regime an observation belongs to, follows a Markov process, which allows us to estimate the transition probabilities to different states/regimes and determine which transitions are most likely to occur. This approach employs hidden Markov models where the data generating process depends on transitions among states, which we can specify

⁸Of course, if we have reliable prior information about where the breaks occur, we should incorporate it. Even if one structural break is appropriate for a given data set, the Chow test depends on restrictive assumptions that can severely compromise its performance (Lax and Rader 2010).

as different periods in longitudinal data.

Park (2010) expands on the hidden Markov model approach, treating the changepoint analysis as a problem of model selection. Each state is represented by a different model with potentially different parameter values. Parameters are allowed to vary across states, and the degree of parameter heterogeneity is essentially determined by comparisons of various non-nested models implied by the states. Transdimensional Markov chain Monte Carlo (MCMC) methods are used to more efficiently explore the model space and parameter values within specific models.⁹

Bai and Perron (1998, 2003) develop a frequentist method for continuous variables that enables us to allow for multiple break points without specifying their locations. It identifies potential break points and provides tests to determine how many breaks, if any, occur in the series. Additionally, it calculates confidence intervals to provide measures of uncertainty around break point estimates. Bai and Perron’s algorithm is based on a dynamic programming approach that computes the optimal number of break points by determining the global minimizers of the residual sum of squares. The sequential procedure involves breaking the series into smaller and smaller possible partitions and checking to see which of these give the optimal fit to the data. The method can be used to assess whether there are breaks in the series itself or in terms of the effects of coefficients. For the former, the approach assesses whether there is a break in the mean of the series, by regressing the time series variable on a constant and then checking whether the intercept varies over the series.

As the discussion in this section indicates, historical researchers have a variety of methods available that enable them to incorporate parameter heterogeneity in ways that capture temporality, periodicity, specificity, and context. All of the methods discussed above can be implemented using software that is publicly available.¹⁰ In the next section, we demonstrate the feasibility of these methodological recommendations with examples of applications drawn

⁹These methods are appropriate for continuous, limited, and qualitative dependent variables. Park (2009) uses this approach to develop a changepoint model for unobserved unit heterogeneity in panel/time-series cross-section data.

¹⁰`WinBUGS` can be used to estimate models with a variety of priors, although it requires that code be written for the priors. `BayesX` enables easy estimation of STAR models with a variety of complex priors (Brezger, Kneib, and Lang 2005). Bai and Perron’s method can be estimated with publicly available `GAUSS` code or using the `strucchange` package for R.

from recent work.

4 Demonstrations

Wawro and Schickler’s (2006) analysis of coalition sizes is a study where the kinds of methods we are advocating could prove useful. In their analysis, they focus on the impact of the adoption of a supermajority cloture rule in the Senate in 1917. They find that coalition sizes on significant legislation generally increased after the adoption of cloture and conduct a number of tests to explain why this is the case. Their basic story is that norms and other constraints against parliamentary obstruction began to break down around the turn of the century, which led senators to seek a formal rule to curtail the use of filibusters. Contrary to what others have claimed, they contend that cloture was a meaningful reform that helps to explain the increase in coalition sizes by providing an institutional mechanism that senators could invoke to reduce the uncertainty surrounding the passage of legislation by building larger (i.e., supermajority) coalitions.

Throughout most of their analysis, they assume that it is valid to break their time series on coalition sizes on significant legislation (which covers 1881 to 1946) into two parts—before and after the adoption of cloture in the 65th Congress (1917–1919). Yet it could be the case that the periodization of the series is more complicated, involving more structural breaks located at different time points. If there are more structural breaks in the data and if they are located far away from when the reform took place, this would call into question their conclusions about coalition sizes and their relation to cloture reform.

This is a tailor-made application for the Bai and Perron method discussed above. To implement this test, we use the average coalition size in each congress covered by the data as the dependent variable and allow for up to five breaks in the series.¹¹ The sequential procedure for selecting breaks indicates that indeed only one break exists in the data, and that occurs at the 63rd Congress (1913–1915). The 90 percent confidence interval for this break indicates that it may also have happened as early as the 57th Congress (1901–1903) or as late as the 69th (1925–1927).¹² This is consistent with Schickler and Wawro’s argument

¹¹We cannot use individual bills since most congresses passed more than one significant bill.

¹²These numbers were produced from GAUSS code written by Perron.

that the dynamics of obstruction had changed during this period, but suggests a more nuanced story. Members of the Senate appear to have begun to build larger coalitions prior to the adoption of cloture, possibly attempting to thwart obstructionists by denying them the resources that a larger opposition coalition would possess. As senators saw the success that larger coalitions brought, it would make formal institutional reinforcement of this strategy all the more attractive. No additional breaks occur through the end of the time series, indicating that the trend in larger coalitions did not reverse or change in significant ways. The method employed allowed for more complex periodization in the data, but none emerged.

We also replicated part of Cox and McCubbins's (2005) analysis examining the ability of the majority party to use agenda setting to move policy away from the median of the legislature as a whole, toward the median of the majority party. They focus on the importance of the Reed Rules, which were initially adopted in 1890 to crack down on parliamentary obstruction in the U.S. House of Representatives. In Chapter 4 of *Setting The Agenda*, they investigate the impact of the Reed Rules in moving policy outcomes toward the majority party's position and find a major impact for a dummy variable indicating whether or not a Congress occurs in the post-Reed era or not. We applied the Bai-Perron method to the dependent variable that Cox and McCubbins use in their analysis—the proportion of final passage bills that move policy toward the majority. The sequential selection procedure indicates a break at the 53rd Congress. Although the Reed Rules were initially adopted in the 50th Congress, they were repealed when the Democrats assumed control in the 52nd and not fully reinstated until the 54th. Thus, this estimated break point makes sense. Interestingly, the 90 percent confidence interval for this break indicates that it could have occurred as early as the 51st Congress or as late as the 58th. The estimated location of the break and the width of this interval may be due to the fitful adoption of the rules.

Farhang and Katznelson's (2005) investigation into sectional influences on the construction of labor policy in the New Deal and Fair Deal eras offers a more involved replication for demonstrating the usefulness of the methods that we advocate. A key claim of Farhang and Katznelson is that important changes in Democratic support for labor-friendly policies took place at this critical time in American history. In the early years of the New Deal, southern Democrats behaved more like their northern colleagues because labor policy was explicitly de-

signed so that it would not interfere with the southern system of racial apartheid. Specifically, domestic and agricultural sectors—occupations in which the majority of African-Americans were employed—were largely exempted from New Deal labor protections, contributing to a long-term set of Faustian bargains where liberals outside the South allowed a distinctly illiberal social and political order to perpetuate for the sake of securing and maintaining national majority coalitions. However, as labor unions began to make inroads in the South, senators and representatives from that section became less willing partners in the New Deal coalition. Of particular concern were labor mobilization efforts in the South that threatened to undercut Jim Crow through the partial integration of unions and the challenge that they posed to the region’s racialized, low-wage political economy. In response, southerners became less likely to vote with their northern counterparts. Labor questions provided the main venue for the emergence of the Conservative Coalition that linked Republicans and southern Democrats in votes that resisted a more robust federal government.

Farhang and Katznelson analyze roll call voting data from the 73rd through the 80th Congress (1933–1948), reporting “likeness scores” that clearly demonstrate the changing coalitional patterns with regard to votes that concerned unions and labor markets. The case for the role that unionization plays is primarily made through narratives that discuss in detail particular cases of legislation considered during this period. The methods that we have advocated in this paper offer an alternative and complimentary approach to test the unionization hypotheses. The story that Farhang and Katznelson tell can be conceived as one of changing parameters. It is not simply just that—as union activity in their region increased—southerners became less likely to support pro-labor policies. The southern reaction to union efforts became more intense over time as they perceived a more severe threat to Jim Crow. Different kinds of parameter variation would capture more precisely how southern attitudes changed. The following multilevel voting model would be appropriate for testing the unionization hypotheses:

$$\Pr(y_{ijkm} = 1) = \text{logit}^{-1}(\alpha_{j[i],k[i]} + \beta_{j[i],k[i]}u_{ij} + \delta' \mathbf{x}_{ijm}) \quad (1)$$

where $\Pr(y_{ijkm} = 1)$ is the probability that senator i from region k in Congress/period j votes the pro-labor position on rollcall m . The subscripts indicate the various levels where

we think there is potentially interesting variation. The variable u_i is a measure of union activity and \mathbf{x}_{ij} represents other measurable factors that affect the vote choice, but do so in a way that is constant across regional and temporal dimensions.

The parameters α and β are allowed to vary over region and time period.¹³ This captures an important nuance in Farhang and Katznelson’s argument. First, there is a temporal component: it could be southerners not only responded negatively to union activity, but that this response became more intense over time as it was perceived as a greater threat. Thus, there is a periodicity to the marginal effect of union activity. Second, the circumstances surrounding a particular vote or set of votes might produce contextual effects. Union efforts during World War Two, for example, under conditions of a tight labor and an ideological campaign against Nazi racism, may have been more menacing to southern representatives than equivalent activities before the war. Finding differences in parameter estimates across levels j and k indicates that context and periodicity were an important part of the unionization hypothesis, and that these innovations to the model are worthwhile.

We adopt undirected autoregressive priors for $\alpha_{j,k}$ and $\beta_{j,k}$. The priors suggest a smooth evolution of the parameters over time, explicitly tying together parameter values in a given period with the values in the two previous and two future periods, and implicitly tying together parameters in non-adjacent periods.¹⁴

To measure union activity, we collected data on work stoppages published by the Bureau of Labor Statistics (BLS) for this period. We weighted this variable with state population, linearly interpolated to fill in the years between the decennial census.¹⁵ This measure of

¹³Following Farhang and Katznelson, we designate the South as constituted by the seventeen states that mandated racial segregation before the 1954 Brown decision. In turn, the South is divided into the seven Deep South states that first seceded from the Union in 1860 (Alabama, Florida, Georgia, Louisiana, Mississippi, South Carolina, and Texas) and the Border South of the other four who left the Union as well as the other six (Arkansas, Delaware, Kentucky, Maryland, Missouri, North Carolina, Oklahoma, Tennessee, Virginia, and West Virginia, which then was part of Virginia). The country’s other states are designated as Non-South.

¹⁴While we are not completely comfortable with some of the independence assumptions that the multilevel modeling approach makes with respect to random effects and explanatory variables, we are for the moment willing to suspend disbelief to see what improvements this approach can offer.

¹⁵We are somewhat concerned about the consistency of this data over the period of investigation, since the BLS appears to have compiled this data from state departments of labor, and there are likely variations in methods of collection. Still, this seems to be the best data available for conducting a quantitative analysis

labor activity enables us to fully explore potential variation that we expect to see across the New Deal and Fair Deal periods. Farhang and Katznelson’s aggregate analysis suggests that the unity that existed between southern and non-southern Democrats in the early New Deal had already begun to break down in the 75th Congress. Thus, including earlier congresses is essential if we are to tease out changes in the parameters. Since we expect important differences across parties, we include a Democratic party indicator in \mathbf{x}_{ij} .

We estimated the model in Eq. 1 using Markov chain Monte Carlo methods. We estimated the model using an expanded set of labor-relevant roll call votes in order to bolster generalizations and obtain a more complete picture of the evolution of roll call voting behavior over time.¹⁶

For purposes of comparison, we also estimated a set of congress-by-congress regressions imposing no priors on the relationship among parameters. This gives us a baseline for comparing the performance of the more sophisticated approach that we advocate. The simple logit results are reported in Table 1 and Figure 1. The coefficients jump around a fair amount and several of them are estimated very imprecisely—of particular note is the wide credible intervals for the estimates of the coefficients on labor activity for Deep South senators in the 76th Congress and Border South senators in the 73rd. Although the results hint at some possible trends over time in the parameters, the imprecision of the estimates may be masking them.¹⁷

The results for the model that uses undirected autoregressive priors are reported in Table

of the scale and detail that we think is essential to make progress on the question. While Troy and Sheffin’s (1985) state-level data on the percentage of the non-agricultural workforce organized in unions is perhaps a more direct measure of union activity, the data is available only for 1939 and 1953. Filling in the intervening years by linear interpolation is likely to introduce serious measurement error, since this would fail to capture nonlinearities in unionization time series across this period. National-level data on unionization indicates that it peaked prior to the passage of the Taft-Hartley Act in 1947 and then declined thereafter.

¹⁶Farhang and Katznelson’s original data set did not include all Congresses from the 73rd to the 80th. We ran three chains in parallel and used the Gelman-Rubin statistic to assess convergence. The model required 500,000 simulations to converge, which while large, is not surprising given the nature and degree of parameter variation in the model. More details about model convergence are available from the authors.

¹⁷We do not report the separate estimates of the coefficient on the Democrat dummy here because we are more interested in the variation over time in the intercepts and in the coefficient on the measure of union activity.

2 and Figures 2 and 3 (the temporal patterns are more easily discernible in the figures; cf. Kastle and Leoni 2007). The parameter estimates are much more stable over time and display more precision compared with the congress-by-congress regressions (note that the scale of the variation in the parameters is much smaller with the Bayesian approach). The use of priors has helped to tease out interesting patterns of variation over time in both the α and β parameters. The α estimates indicate how the baseline probability of voting in the pro-labor direction changes across region and congresses. For Deep South and Border South senators, this baseline probability decreases as we progress from the New Deal to the Fair Deal periods. There is a clear distinction in the α estimates—that is, no overlap in the credible intervals—between the 73rd and 74th Congresses and the 78th and 79th Congresses for Deep South senators. A similar pattern is seen for Border South senators, with clear separation from the earlier congresses in the period occurring in the 79th and 80th Congresses. The Non-South α estimates pertain to the baseline voting patterns of Republicans (i.e., non-southerners who are not Democrats) and do not show a clear pattern over time. While they first appear to decrease and then increase, the overlap among the credible intervals suggest that there is not much of a difference in these estimates over the period in questions. These results are consistent with Farhang and Katznelson’s aggregate level analysis of coalition patterns. The temporal pattern indicates clear separation in the baseline voting patterns of southern Democrats, with the former generically becoming more anti-labor.

The β coefficients capturing the variable effects of labor unrest also display interesting variation over time that reveals a fascinating subordinate trend inside the larger pattern identified by Farhang and Katznelson. Within the aggregate pattern of a declining inclination of southern Democrats to vote in support of labor, we discover that constituency labor action produced a greater inclination for pro-labor voting. Southern members, these results suggest, were cross-pressured, caught between their overarching desire to protect a system of white supremacy and their need to be responsive to their voters, some of whom were drawn to unions and their capacity to improve life situations. We can see this pattern developing over time. For Deep South senators, the coefficients are not bounded away from zero for the 73rd through 77th Congresses, suggesting no relationship between union activity and

voting on labor-related legislation. But the coefficients in the 78th–80th Congresses are bounded away from zero and positive, indicating an increase in the likelihood of voting the pro-labor position. For Border South senators, the coefficients estimates are statistically different from zero and positive for the 77th–80th Congresses. For non-southerners, increases in work stoppages are associated with a decrease in pro-labor voting in the 73rd and 80th Congresses, but indicate an increase in labor support in the 75th. The β estimates are positive and bounded away from zero for border state senators in the 78th, 79th, and 80th Congresses and for Deep South senators in the 80th. The coefficient on the Democrat dummy indicates, not surprisingly, that Democrats are more likely to vote pro-labor than are Republicans, all else equal. Despite a general southern propensity to defect from the median party position on labor, there is clear evidence that union activity in the South is associated with an increase in pro-labor voting in the 1940s. This pattern within a pattern is revealed by the methods we have applied, and was missed not only by Farhang and Katznelson but by all the extant literature on unions and the South written by historians and social scientists.

Turning to the marginal effects, Figure 3 displays differences in in the predicted probability of voting the pro-labor position (with credible intervals to indicate estimation uncertainty) when our measure of union activity is at its median value and then increased by two standard deviations. We see substantively significant increases in the predicted probability of voting the pro-labor position for Deep South and Border South senators in the 79th and 80th Congresses (for other congresses the probability of voting pro-labor are not bounded away from zero). Deep South senators’ likelihood of voting pro-labor increases approximately 16 percentage points in these congresses, while senators from Border South states are approximately 22 percentage points more likely to vote pro-labor. Non-southern Democrats are predicted to experience *decreases* in the likelihood of voting pro-labor in the 73rd and 80th Congresses—about .21 percentage points in the former and .09 percentage points in the latter. In the 75th Congress, non-southern Democrats are about 10 percentage points more likely to vote pro-labor when there is an increase in the union activity measure.

The positive relationship that we see between southern senators’ voting patterns and labor activity in their states reveals an interesting and significant modification to the story of labor policymaking in this period. One possible explanation for the results is that southerners felt

cross-pressured on labor votes. While part of their constituencies did not like the threat to desegregation and the southern order that union activity presented, other constituents were actually being mobilized to participate in work stoppages, which perhaps meant that they could have been more easily mobilized to oppose reelection-seeking senators who were not sympathetic enough to labor. Southerners in Congress may have hedged their bets in places where unions were experiencing the most success at organizing. The larger marginal effect for border state senators suggest that they felt more cross-pressure than those from the Deep South, where preferences for segregation were stronger.

This replication demonstrates the usefulness of including parameter heterogeneity but constraining it using history-relevant priors. Using undirected autoregressive priors enabled us to borrow strength from temporally adjacent congresses in order to improve the precision of parameter estimates and uncover interesting patterns in the evolution of roll call voting behavior. It also suggests a larger potential to disclose historical arrangements, relationships, and processes that otherwise would remain obscure, thus shifting the character of objects of analysis and opening the way for the development of fresh hypotheses to account for what happened within the richer universe of behavior that now can be ascertained.

5 Discussion

This paper has been motivated by the concern that as the turn to historical research proceeds in political science, forms of mutual incomprehension, semi-accurate caricature, and concern about the proper range of assumptions and research methods have erected barriers separating scholars with richly nuanced qualitative approaches to history from colleagues who make quantitative methods central to their studies of political behavior and institutions. If we do not attempt to overcome these barriers by addressing key points of contention, both worlds of scholarship will be less robust than they should be.

Within the framework of identifying the main legitimate reservations about mainstream quantitative scholarship that is historical by those who practice historical research as their primary craft, our main goal is to advocate how thoughtful utilization of a set of approaches that privilege parameter variation can serve as a promising means, though not the only one, to bridge these research communities by systematically capturing just those features about

parameters and their alteration that most interest the historical community. From our perspective, not only is it imperative to take seriously the worry among historians and historical social scientists that quantitatively-oriented political science fails sufficiently to attend to context, historical specificity, temporality, and periodicity, but to acknowledge that unless such matters can be made constitutive features of inquiry, their skepticism will remain justified. Good research on historical periods cannot proceed by a flattening universalism, and certainly not by any such an *a priori* assumption. Rather than expect models to predictably port across time, we should be building models that seek to internalize and reflect central historical features and processes by integrating parameter heterogeneity and complexity inside their very construction.

In turn, though, such efforts, to the extent they succeed by illuminating the historical process, put pressure—a welcome kind of pressure—on political historians and qualitative political scientists to deepen and broaden their research repertoire. It simply is not good enough to express skepticism about, say, mainstream approaches to multivariate analysis and, with varying degrees of dismissiveness and willful unawareness, continue with a “business as usual” attitude that largely ignores advances in empirical modeling and their potential to causally investigate historical processes.

We have discussed specific prescriptions for moving forward to produce more robust historical quantitative analyses that put front and center the concerns that historians have expressed about standard political science methodology. Temporality can be captured by using splines and directed and undirected priors that tap into temporal evolution in parameter effects. Hierarchical modeling is particularly well-suited to incorporate contextual considerations. Structural change or changepoint models, especially given recent advances in their estimation, offer tremendous opportunities to systematically unveil periodicity that can have a profound impact on our evaluation of temporally-centric hypotheses. Beyond this expanded use of existing methods, we also see an imperative to develop new modeling and estimation techniques that address problems unique to historical political analysis.

A long list of fundamental questions could be more thoroughly and compellingly explored through historically-oriented analyses that moved beyond the standard methodological toolkit in the ways that we have advocated. Within American politics, the list includes

how to think about party in relation to preferences; issues concerning the number and substantive meaning of ideological and policy dimensions; questions about the structuration of lawmaking and the impact of organizational features on outcomes; the ways patterns of representation orient constituency ties; the meaning of roll call votes and the balance between those which are partisan and those which are not; and the barriers to legislative productivity, both institutional and behavioral. Even well-tilled areas of inquiry that are inherently historical, such as the debate over the existence and impact of partisan realignments, could be revisited fruitfully by using the modeling innovations we have discussed to more fully resolve the attendant theoretical and empirical controversies. Major questions in other subfields also cry out for more sophisticated methodological approaches to history. Questions of the effects of democracy and trade on peace, of the relationship between economic and democratic development, and of economic and political inequality all have historical dimensions that should be central to the empirical methods used to address them. But inquiries into these questions often proceed not only without attending to the major methodological concerns of those who more routinely conduct historical research, but with assumptions and methods that treat data with uniformity across historical time while discounting variations in time as they might apply to a wide array of relevant causal features.

If we wish to probe these issues, irrespective of our disciplinary orientations and methodological priors, we will need to find ways of working that, at once, are deeply substantive and systematic. Here, we have advocated a set of empirical tools that serve as a means to this end. To the extent that this, among other possibilities, is persuasive, it also implies obligations about the range of issues, literatures, and methods that quite disparate research communities might learn to share, and entails commitments that will be difficult, sometimes painful, to achieve. But the payoff promises to be considerable.

References

- Angrist, Joshua D. and Jorn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press.
- Bai, Jushan and Pierre Perron. 1998. "Estimating and Testing Linear Models with Multiple

- Structural Changes.” *Econometrica* 66 (1): 47–78.
- Bai, Jushan and Pierre Perron. 2003. “Computation and Analysis of Multiple Structural Change Models.” *Journal of Applied Econometrics* 18: 1–22.
- Bartels, Larry. 1996. “Pooling Disparate Observations.” *American Journal of Political Science* 40: 905–42.
- Bates, Robert, Avner Greif, Margaret Levi, Jean-Laurent Rosenthal, and Barry Weingast. 1998. *Analytic Narratives*. Princeton: Princeton University Press.
- Berinsky, Adam J. 2009. *In Time of War: Understanding American Public Opinion from World War II to Iraq*. Chicago: University of Chicago Press.
- Besag, Julian. 1974. “Spatial interaction and the statistical analysis of lattice systems.” *Journal of the Royal Statistical Society, Series B* 36: 192–236.
- Besag, Julian. 1975. “Statistical Analysis of Non-Lattice Data.” *The Statistician* 24: 179–195.
- Besag, Julian, Peter Green, David Higdon, and Kerrie Mengersen. 1995. “Bayesian Computation and Stochastic Systems (with Discussion).” *Statistical Science* 10: 3–66.
- Besag, Julian and Charles Kooperberg. 1995. “On Conditional and Intrinsic Autoregressions.” *Biometrika* 82 (4): 733–746.
- Binder, Sarah A. 1997. *Minority Rights, Majority Rule*. New York: Cambridge University Press.
- Brady, David, Kara Buckley, and Douglas Rivers. 1999. “The Roots of Careerism in the U.S. House of Representatives.” *Legislative Studies Quarterly* XXIV: 489–510.
- Brady, David W. 1973. *Congressional Voting in a Partisan Era: A Study of the McKinley Houses and a Comparison to the Modern House of Representatives*. Lawrence: University of Kansas Press.
- Breslow, N. E. and D. G. Clayton. 1993. “Approximate Inference in Generalized Linear Mixed Models.” *Journal of the American Statistical Association* 88 (421): 9–25.

- Brezger, Andreas, Thomas Kneib, and Stefan Lang. 2005. "BayesX: Analyzing Bayesian Structured Additive Regression Models." *Journal of Statistical Software* 14 (11): 1–22.
- Brezger, Andreas and Stefan Lang. 2006. "Generalized Additive Regression Based on Bayesian P-splines." *Computational Statistics and Data Analysis* 50: 967–991.
- Brunell, Thomas L. and Bernard Grofman. 1998. "Explaining Divided U.S. Senate Delegations, 1788–1996: A Realignment Approach." *American Political Science Review* 92: 391–400.
- Carpenter, Daniel P. 2010. *Reputation and Power: Organizational Image and Pharmaceutical Regulation at the FDA*. Princeton, N.J.: Princeton University Press.
- Chib, Siddhartha. 1998. "Estimation and Comparison of Multiple Change-Point Models." *Journal of Econometrics* 86 (2): 221–241.
- Cooper, Joseph and David W. Brady. 1981. "Toward a Diachronic Analysis of Congress." *American Political Science Review* 75 (December): 988–1006.
- Cox, Gary W. and Mathew D. McCubbins. 2005. *Setting The Agenda: Responsible Party Government in the U.S. House of Representatives*. Cambridge: Cambridge University Press.
- Deaton, Angus. 2009. "Instruments of Development: Randomization in the Tropics, and the Search for the Elusive Keys to Economic Development." NBER Working Paper #14690.
- Fahrmeir, L. and G. Tutz. 2001. *Multivariate Statistical Modelling Based on Generalized Linear Models*. New York: Springer-Verlag.
- Farhang, Sean and Ira Katznelson. 2005. "The Southern Imposition: Congress and Labor in the New Deal and Fair Deal." *Studies in American Political Development* 19: 1–30.
- Gaddis, John Lewis. 2002. *Landscapes of History*. Oxford: Oxford University Press.
- Gelman, Andrew and Jennifer Hill. 2007. *Data Analysis Using Regression and Multi-level/Hierarchical Models*. Cambridge: Cambridge University Press.

- Girosi, Federico and Gary King. 2008. *Demographic Forecasting*. Princeton, NJ: Princeton University Press.
- Hastie, T. J. and R. J. Tibshirani. 1990. *Generalized Additive Models*. London: Chapman and Hall.
- Hofstadter, Richard. 1956. "History and the Social Sciences." In Fritz Stern, editor, *The Varieties of History: From Voltaire to the Present*, New York: Noonday Press.
- Isaac, Larry W. and Larry J. Griffin. 1989. "Ahistoricism in Time Series Analyses of Historical Processes: Critique, Redirection, and Illustrations From US Labor History." *American Sociological Review* 54 (December): 873–890.
- Jenkins, Jeffrey A., Justin Peck, and Vesla M. Weaver. 2010. "Between Reconstructions: Congressional Action on Civil Rights, 1891–1940." *Studies in American Political Development*.
- Kalyvas, Stathis. 1996. *The Rise of Christian Democracy in Europe*. Ithaca and London: Cornell University Press.
- Kalyvas, Stathis. 2006. *The Logic of Violence in Civil War*. New York: Cambridge University Press.
- Kastellec, Jonathan P. and Eduardo L. Leoni. 2007. "Using Graphs Instead of Tables in Political Science." *Perspectives on Politics* 5 (4): 755–771.
- Katznelson, Ira and John S. Lapinski. 2006. "At the Crossroads: Congress and American Political Development." *Perspectives on Politics* 4 (June): 243–260.
- King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry*. Princeton: Princeton University Press.
- Kreuzer, Marcus. 2010. "Historical Knowledge and Quantitative Analysis: The Case of the Origins of Proportional Representation." *American Political Science Review* 104 (2): 369–392.

- Lax, Jeffrey R. and Kelly T. Rader. 2010. "Legal Constraints on Supreme Court Decision Making: Do Jurisprudential Regimes Exist?" *Journal of Politics* 71 (2): 273–284.
- Mansfield, Edward and Jack Snyder. 2005. *Electing to Fight: Why Emerging Democracies Go to War*. Cambridge: MIT Press.
- Moe, Terry M. 1987. "Interests, Institutions, and Positive Theory." *Studies in American Political Development* 2: 236–299.
- Orren, Karen and Stephen Skowronek. 2004. *The Search for American Political Development*. Cambridge: Cambridge University Press.
- Page, Benjamin and Robert Shapiro. 1992. *The Rational Public*. Chicago: University of Chicago Press.
- Park, Jong Hee. 2009. "Joint Modeling of Dynamic and Cross-Sectional Heterogeneity: Introducing Hidden Markov Panel Models." Paper presented at the Summer Meeting of the Society for Political Methodology.
- Park, Jong Hee. 2010. "Structural Change in the U.S. Presidents' Use of Force Abroad." *American Journal of Political Science* 54 (3): 766–782.
- Pierson, Paul. 2004. *Politics in Time: History, Institutions, and Social Analysis*. Princeton: Princeton University Press.
- Pierson, Paul and Theda Skocpol. 2002. "Historical Institutionalism in Contemporary Political Science." In Ira Katznelson and Helen V. Milner, editors, *Political Science: State of the Discipline*, New York: W.W. Norton. Pages 693–721.
- Schickler, Eric. 2000. "Institutional Change in the House of Representatives, 1867–1998: A Test of Partisan and Ideological Power Balance Models." *American Political Science Review* 94: 269–287.
- Sewell, William H. 2005. *Logics of History*. Chicago: Chicago University Press.
- Silbey, Joel H. 2000. "Current Historiographic Trends in the Study of the Twentieth-century Congress." *Social Science History* 24: 317–331.

- Tilly, Charles and Robert E. Goodin. 2006. "It Depends." In Robert E. Goodin and Charles Tilly, editors, *The Oxford Handbook of Contextual Political Analysis*, New York: Oxford University Press.
- Trachtenberg, Marc. 2006. *The Craft of International History: A Guide to Method*. Princeton: Princeton University Press.
- Troy, Leo and Neil Sheffin. 1985. *U.S. Union Sourcebook*. West Orange, NJ: Industrial Relations Data and Information Services.
- Wawro, Gregory and Eric Schickler. 2006. *Filibuster: Obstruction and Lawmaking in the United States Senate*. Princeton: Princeton University Press.
- Western, Bruce and Simon Jackman. 1994. "Bayesian Inference in Comparative Research." *American Political Science Review* 88: 412–423.
- Western, Bruce and Meredith Kleykamp. 2004. "A Bayesian Change Point Model for Historical Time Series Analysis." *Political Analysis* 12 (4): 354–74.

Table 1: Results for α and β parameters from congress-by-congress logistic regression analysis of labor roll call votes in the Senate (73rd–80th Congresses)

	$\hat{\alpha}$	$\hat{\alpha}_{SE}$	$\hat{\beta}$	$\hat{\beta}_{SE}$
73rd,Deep South	0.90	0.33	0.15	1.00
74th,Deep South	0.85	0.47	0.70	0.91
75th,Deep South	-0.99	0.25	2.17	0.74
76th,Deep South	0.47	0.55	-0.03	2.71
77th,Deep South	0.14	0.56	-0.04	0.31
78th,Deep South	-1.29	0.39	0.62	0.37
79th,Deep South	-0.71	0.18	0.16	0.16
80th,Deep South	-0.73	0.19	1.09	0.23
73rd,Border South	0.58	0.31	-4.29	1.51
74th,Border South	-0.05	0.28	0.02	0.19
75th,Border South	0.31	0.29	-0.39	0.42
76th,Border South	-0.20	0.39	0.46	0.30
77th,Border South	-0.08	0.52	0.29	0.22
78th,Border South	-0.83	0.25	0.08	0.11
79th,Border South	-0.68	0.12	0.16	0.04
80th,Border South	-1.51	0.15	0.32	0.06
73rd,Non-South	-0.29	0.16	-0.70	0.15
74th,Non-South	-0.71	0.29	-0.09	0.26
75th,Non-South	-1.43	0.19	0.18	0.06
76th,Non-South	-1.06	0.39	1.19	0.40
77th,Non-South	-2.42	0.64	-0.06	0.19
78th,Non-South	-0.93	0.19	-0.09	0.09
79th,Non-South	-0.99	0.09	-0.04	0.03
80th,Non-South	-1.28	0.14	-0.28	0.10

Notes: Table entries are logit estimates obtained by maximum likelihood. “SE” indicates estimated standard errors for the relevant parameters.

Table 2: Results for α and β parameters from the analysis of labor roll call votes in the Senate using undirected autoregressive priors (73rd–80th Congresses)

	$\hat{\alpha}$	$\hat{\alpha}_{SE}$	$\hat{\beta}$	$\hat{\beta}_{SE}$
73rd,Deep South	1.15	0.30	0.50	0.56
74th,Deep South	0.77	0.27	0.51	0.45
75th,Deep South	-0.04	0.20	0.49	0.36
76th,Deep South	0.11	0.33	0.44	0.27
77th,Deep South	-0.29	0.32	0.42	0.23
78th,Deep South	-0.76	0.20	0.44	0.19
79th,Deep South	-0.65	0.12	0.48	0.14
80th,Deep South	-0.06	0.12	0.57	0.22
73rd,Border South	0.14	0.36	0.05	0.22
74th,Border South	0.09	0.23	0.07	0.17
75th,Border South	0.06	0.15	0.09	0.13
76th,Border South	-0.04	0.15	0.11	0.10
77th,Border South	-0.20	0.14	0.14	0.07
78th,Border South	-0.40	0.10	0.17	0.04
79th,Border South	-0.61	0.07	0.20	0.03
80th,Border South	-0.87	0.11	0.25	0.04
73rd,Non-South	-1.08	0.22	-0.60	0.17
74th,Non-South	-1.52	0.16	-0.18	0.12
75th,Non-South	-1.78	0.11	0.17	0.06
76th,Non-South	-1.12	0.30	0.19	0.12
77th,Non-South	-1.91	0.35	0.09	0.13
78th,Non-South	-1.39	0.12	-0.00	0.08
79th,Non-South	-1.15	0.07	-0.04	0.03
80th,Non-South	-1.21	0.07	-0.21	0.07
Democrat			2.47	0.07

Notes: Estimation using MCMC implemented in `WinBugs` employing 3 chains, each with 500,000 iterations (first 250,000 discarded), 1,000 iterations saved. “SE” indicates estimated standard errors for the relevant parameters.

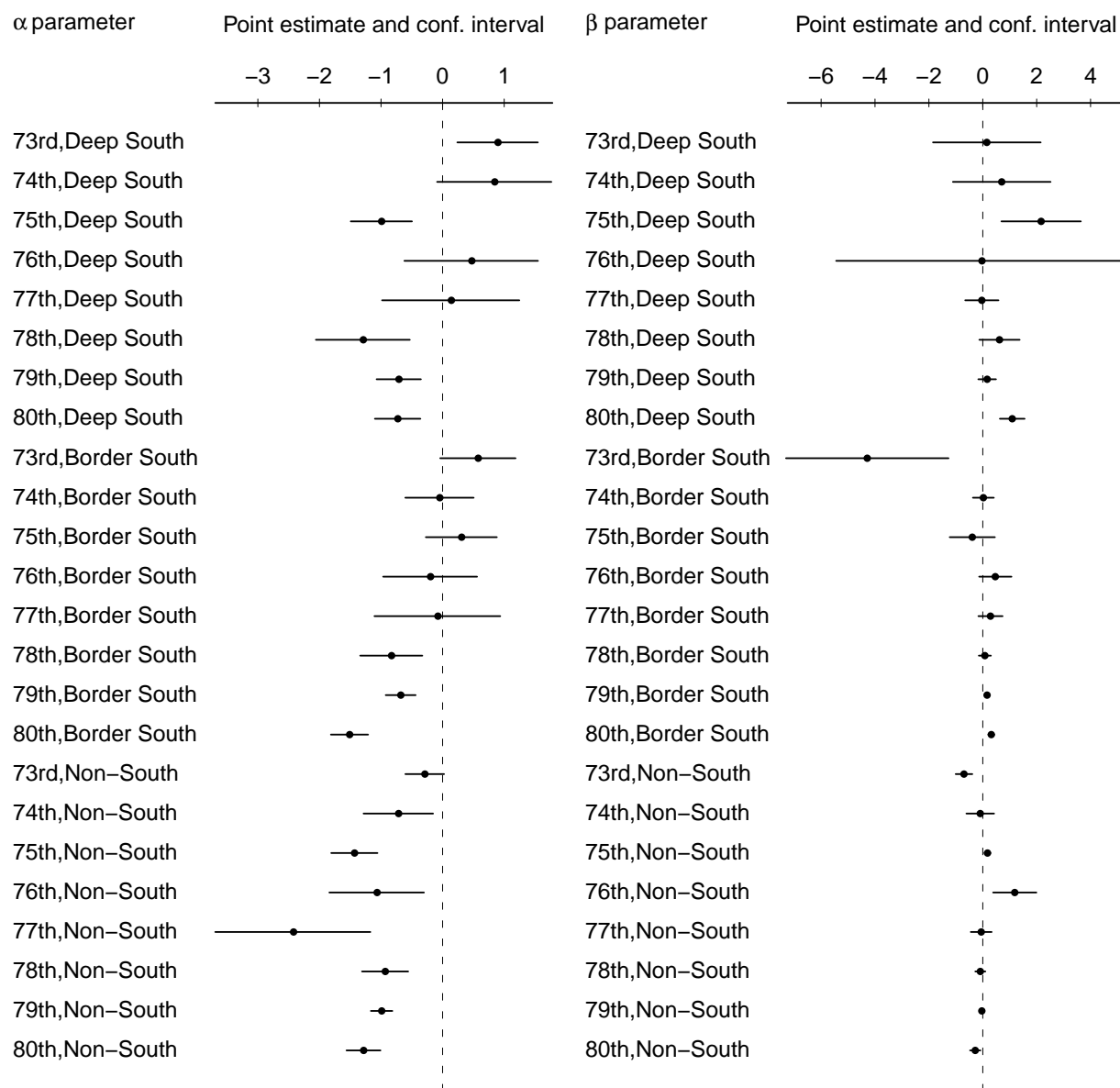
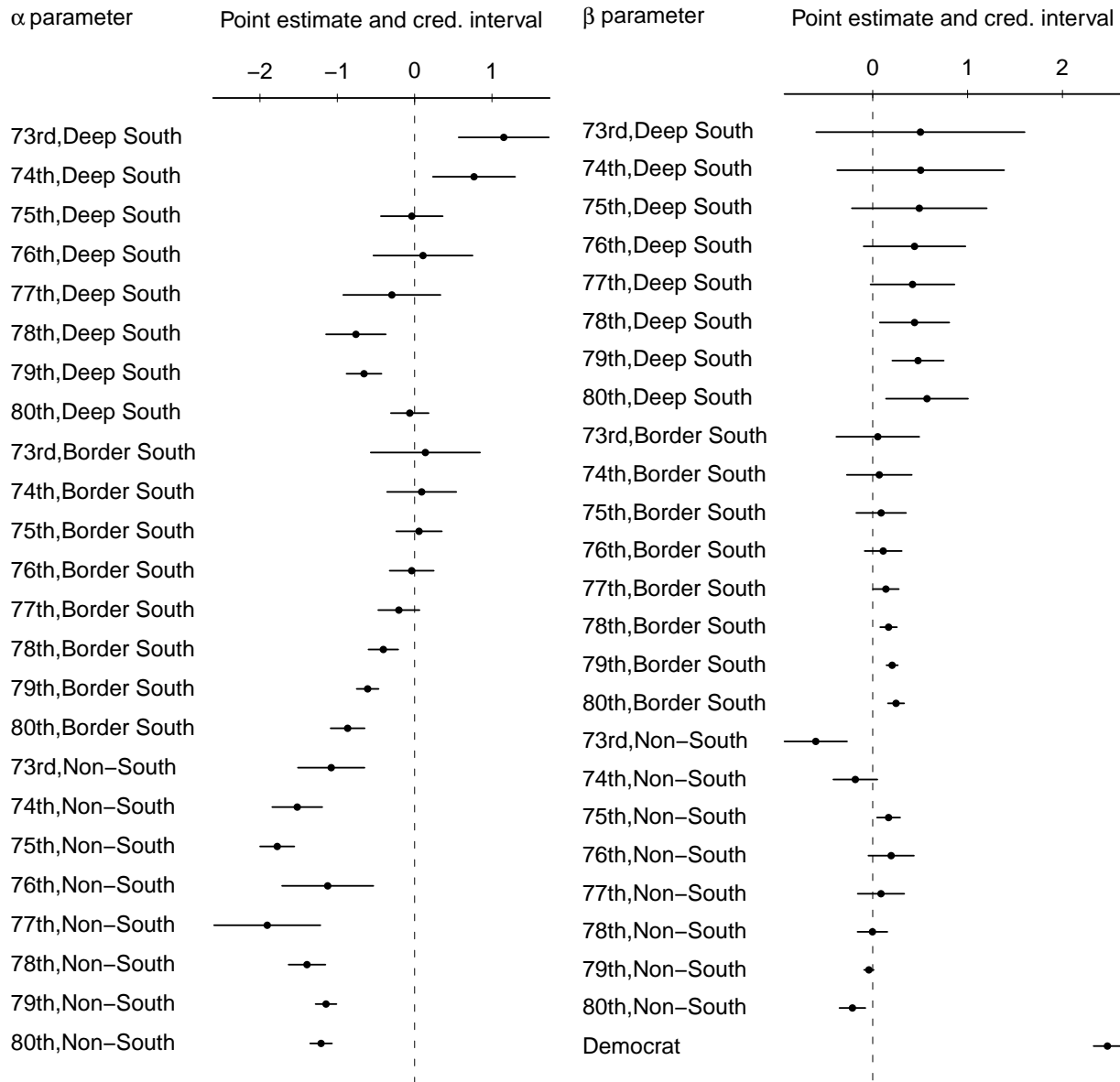


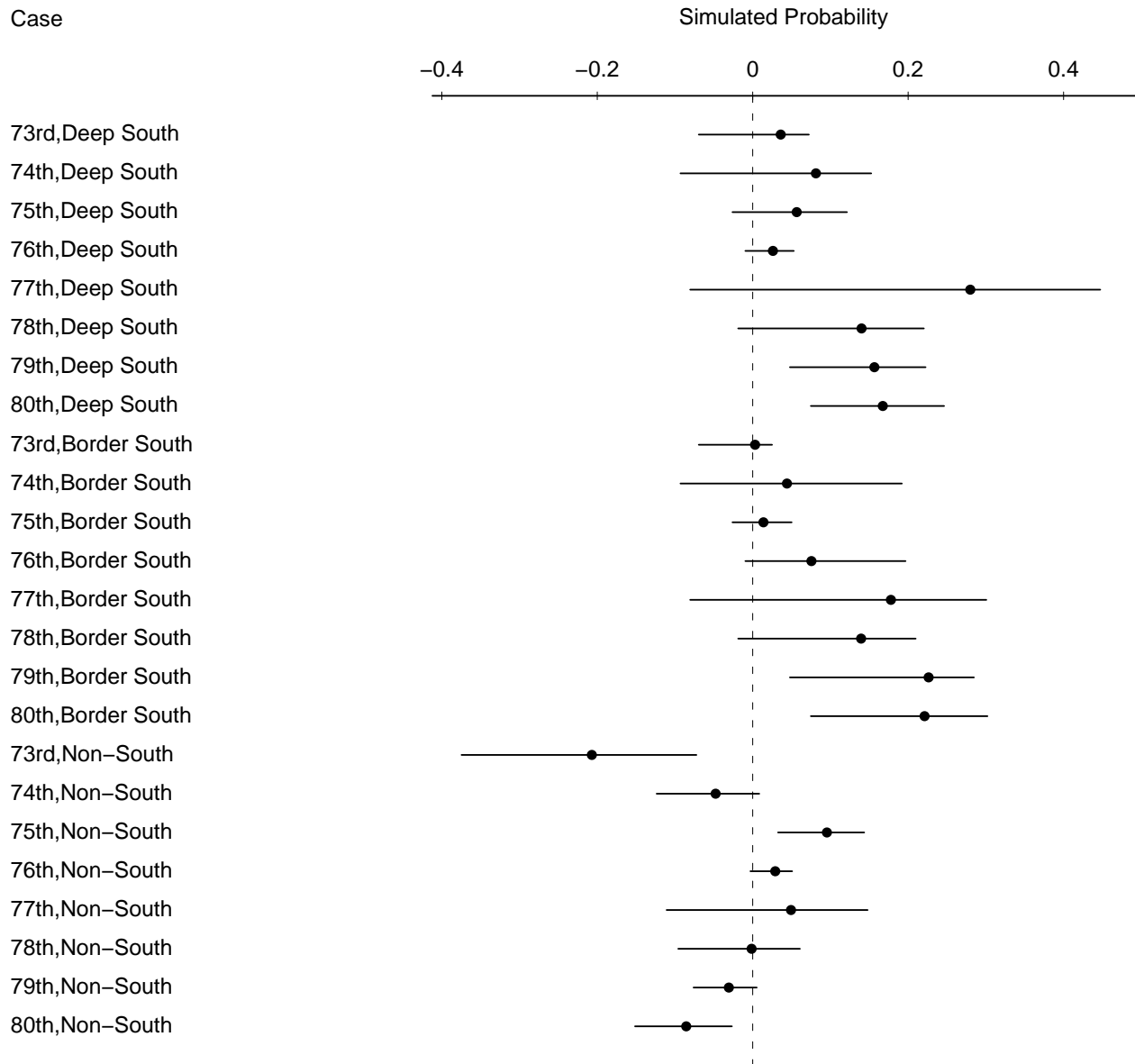
Figure 1: Estimated coefficients and 95% confidence intervals for α and β parameters from congress-by-congress logistic regression analysis of labor roll call votes in the Senate (73rd–80th Congresses).

Figure 2: Estimated coefficients and 95% credible intervals for α and β parameters from the analysis of labor roll call votes in the Senate using undirected autoregressive priors (73rd–80th Congresses).



Note: Estimation using MCMC implemented in WinBugs employing 3 chains, each with 500,000 iterations (first 250,000 discarded), 1,000 iterations saved.

Figure 3: Simulated probabilities (with 95% credible intervals) for Democratic senators



Note: The circles represent the differences in the simulated probabilities when the variable values are set to median values for the region and period and when the work stoppages variables is increased by two standard deviations. The lines represent 95% credible intervals.