

A Dynamic Dirichlet Process Mixture Model for the Partisan Realignment of Civil Rights Issues in the U.S. House of Representatives*

Nuannuan Xiang[†]

Yuki Shiraito^{‡§}

This draft: October 14, 2024

First draft: July 7, 2020

Abstract

The positions of the two major parties in the United States on civil rights issues reversed in the 20th century. The conventional wisdom views the reversal as a structural break in the 1960s led by party elites, whereas recent work argues that the change occurred gradually from the 1930s driven by local activists within both parties. To address this debate, we develop a nonparametric Bayesian model that incorporates the hidden Markov model into the Dirichlet process mixture model. In analyzing rank-and-file legislators' behavior, we model the emergence and disappearance of their latent voting blocks as a gradual process rather than a one-time structural change, thereby identifying both steady and sudden changes of voting coalitions. Our analysis shows that, in addition to gradual changes of party positions beginning in the 1930s, Democrats and Republicans in southern States formed a voting coalition from 1950s.

*We thank Dean Knox, Christopher Lucas, Jacob Montgomery, Iain Osgood, Mark Pickup, Kevin Quinn, and audiences at the 2020 Winter Meeting of Japanese Society for Quantitative Political Science, the 37th Annual Summer Meeting of the Society for Political Methodology, and Virtual Speaker Series in the Department of Political Science at Washington University in St. Louis for insightful comments and suggestions.

[†]Assistant Professor of Health Policy and Management, Mailman School of Public Health, Columbia University. Email: nx2174@cumc.columbia.edu.

[‡]Assistant Professor, Department of Political Science, University of Michigan. URL: shiraito.github.io.

[§]Correspondence: Center for Political Studies, 4259 Institute for Social Research, 426 Thompson Street, Ann Arbor, MI 48104-2321. Phone: 734-615-5165, Email: shiraito@umich.edu.

1 Introduction

Gradual structural changes—changes that are obscure while they are happening, but that are obvious after a few decades—often pose debates among social scientists. Since the process of changes is not directly observable, researchers need to infer when the latent shift began and how it evolved. Political scientists studying party politics in the United States are well aware that the Democrats and the Republicans switched their policy positions on racial issues in the 20th century. Today, the Democratic Party is the advocate for racial liberalism, whereas the Republican Party is associated with racial conservatism. However, the party positions were the exact opposite in the 1930s. The conventional wisdom (e.g., Garmines and Stimson, 1989; McAdam and Kloos, 2014; Edsall and Edsall, 1992; Califano, 1992) has viewed this change as a structural break, whose high point was marked by civil rights legislation in the 1960s, driven by party elites in Washington D.C. On the other hand, Schickler (2016) challenges this view by arguing that rank-and-file legislators’ behavior began altering in the 1930s. These arguments attempt to identify the timing and development of a latent change underlying observable actions in the Congress.

To empirically address this debate, we develop a dynamic model of gradual structural changes that extends the Dirichlet process (DP) mixture model (Ferguson 1973; Antoniak 1974; see also Teh 2010). The DP mixture model (DPMM) is a nonparametric Bayesian model for clustering units into latent groups. It places the DP prior on the mixing distribution of a mixture model, thereby allowing the number of clusters to be estimated from data, instead of requiring it to be specified *ex ante*. We extend the DPMM to a dynamic setting by modifying the standard DP prior to a Markov process of the DP for cluster membership assignments over time. Similarly to the static DP, this dynamic DP is closely connected to the Chinese restaurant process (CRP; Blackwell and MacQueen 1973). We call it the intergenerational Chinese restaurant process (IgCRP), because the cluster assignments in each time period follows the CRP conditional on the cluster assignments of all units in the previous period. A distinctive feature of our model is that it models a process in which multiple clusters emerge and diminish as a continuing process rather than a structural break. At the same time, our model retains the popular feature of the DPMM that the number of clusters does not need to be specified *ex ante*. Moreover, as is the case with the DPMM, a broad class of models can be the mixture component of the dynamic DPMM, and therefore it is widely applicable beyond our specific application to the party position switching on racial issues.

We analyze the evolution of party positions on civil rights issues in the U.S. House of Representatives using the dynamic DPMM. The data set consists of various legislative actions measured repeatedly for each House member from the 73rd to the 92nd Congress (1933–1973). Our analysis builds on the voting block model of Spirling and Quinn (2010), which uses the DPMM to identify latent clusters of legislators based on their voting behavior. However, our dynamic DPMM allows us to make inference on gradual changes in the cluster memberships of legislators over time, thereby answering the question of when and how the parties switched their positions on racial issues. We find that the Democratic legislators from the North began moving toward pro-civil rights

positions in the 1930s and that the southern Republicans formed an anti-civil rights coalition with the southern Democrats in the 1950s. These findings are consistent with Schickler (2016), who argue that the party position switching on racial issues began in the 1930s and that the northern Democrats were the key players in the process.

Our extension of the DPMM is characterized by clusters shared across time and dependence of cluster assignments on the previous period. The methodological challenge of our application is to capture gradual changes in latent heterogeneity over time. We use a mixture model to account for latent heterogeneity, which is common in political science research (Imai and Tingley, 2012; Spirling and Quinn, 2010; Kyung et al., 2011). However, the change-point modeling approach, which is also common to model temporal dynamics in social sciences (Park, 2010; Pang et al., 2012; Kim et al., 2020) is not suitable due to its assumption that all units switch their status simultaneously and hence latent clusters are not shared across time. We develop a new dynamic DP prior that combines temporal dependence similar to Caron et al. (2012) and stickiness of a dynamic DP introduced by Fox et al. (2011).

The remainder of the paper proceeds as follows. The next section introduces our motivating empirical example, the racial switch of the Democratic Party and the Republican Party. Section 3 first describes the proposed model using the stick-breaking definition of the Dirichlet Process. Then the section illustrates the intergenerational Chinese restaurant metaphor for the proposed model, which provides the intuition on how latent clusters evolve over time. Section 4 presents the empirical analysis of the motivating example, followed by concluding remarks.

2 Motivating Example and Data

2.1 The Racial Realignment, 1932-1965

Our motivating example is the shifting party coalitions for civil rights in the US from the 1930s to the 1970s analyzed by Eric Schickler’s book of *Racial Realignment: The Transformation of American Liberalism, 1932-1965* (Schickler, 2016). In his book, Schickler argues that the Democratic and Republican parties’ positions on civil rights started changing in the 1930s, much earlier than the conventional wisdom.

The two parties in the US switched their positions on racial issues in the 20th century. Today, the Democratic Party is associated with racial liberalism, supporting government efforts to redress racial inequality, while the Republican party is associated with racial conservatism, resisting governmental interventions in racial issues. Before the 1930s, however, the two parties’ positions on racial issues were the reverse. The Republican Party used to be more identified with African Americans than the Democratic Party.

Scholars often view the reversal in the two parties’ positions as a sudden structural break that occurred in the 1960s. According to this view, local Republicans were slightly more liberal on racial issues than their Democratic counterparts in the 1940s and 1950s (Garmines and Stimson, 1989). Although northern Democrats had an incentive to support civil rights to seek support from

African American as increasingly important voters, federalism prevented local politicians from making commitment to programmatic liberalism (Weir, 2005). Instead, party elites in Washington, D.C. led the change (McAdam and Kloos, 2014; Edsall and Edsall, 1992). The “critical juncture” arrived during the presidential election of 1964, when Democratic candidate Lyndon B. Johnson and Republican candidate Barry Goldwater took sharply different positions on civil rights issues (Califano, 1992). Local party activists then followed national leaders to change their racial positions.

In contrast to the structural-break view of the reversal, Schickler (2016) makes a critical observation that the change of parties’ racial positions was a gradual process that started in the mid-1930s. Rather than national party elites, locally oriented rank-and-file party members drove the gradual change. Schickler’s view aligns with recent scholarship that challenges the traditional insight of a sudden shift of the two parties’ positions on civil rights (Farhang and Katznelson, n.d.; Chen et al., 2008).

Specifically, Schickler (2016) emphasizes that northern Democrats at the local level took the lead in making the change.¹ Local northern Democrats were gradually transformed by the New Deal coalition with the Congress of Industrial Organizations (CIO), African Americans, and other urban liberals. Initially, the New Deal coalition, which was made possible by the shared interests of economic liberalism, had little to do with race. As Democrats’ nonpartisan allies increased their civil rights advocacy, so were northern Democrats. While national party elites had a strong interest in maintaining the solidarity of the party, pressures from local party activists forced elites in Washington, D.C. to break the traditional North-South coalition. At the same time, the Republican party, which used to be more supportive of civil rights, was gradually divided over racial issues, as the demand for pursuing civil rights in some of its constituencies waned.

2.2 Data

Schickler (2016) provides rich data to assess his argument, including historical survey data to trace the configuration of economic and racial liberalism at the macro level, data of state party platforms to understand how state Democratic and Republican parties positioned themselves on civil rights, and data concerning congressional action on civil rights to measure House members’ positions on racial issues. In this paper, we reanalyze the congressional data to understand how House members made coalitions across states and party lines, and how the coalitions changed gradually over the years.

The data of congressional action includes roll-call votes, signing discharge petitions to advance civil rights bills, floor speech, as well as bill sponsorship.² Summary statistics of the data are shown in Supplementary Information (SI) A.

¹Following Schickler (2016), we define “southern” states or the “South” as the 11 Confederate states plus Kentucky and Oklahoma. “Northern” states or the “North” refers to all other states.

²Roll-call votes are most commonly used to measure members’ preferences. However, Schickler points out that roll calls could be a misleading measure. Only a very small portion of civil rights bills reached the floor. Most civil rights bills were blocked by the House Rules Committee, which was dominated by southern senior Democrats.

Roll-call Votes. Following the original work of Schickler (2016), we use a dummy variable to indicate whether a House member supported a civil rights bill or not. “1” means that a member voted “yes” for a bill, and “0” indicates that the member either voted “no” or was absent. In Table A.1, we list the number of roll calls about civil rights from the 73rd Congresses (1933-35) to the 92nd Congress (1971-73). It shows that during the period of this study, not every Congress had roll calls of civil rights issues. The table also presents the means and standard deviations of the roll-call variable by parties and years. The standard deviations are as large as the means, meaning the within-party heterogeneity is high.

Discharge Petition Signatures. If a legislative committee or the Rules Committee has blocked a bill from reaching the floor for a long time, House members can sign a petition to ask for advancing the bill.³ We use a dummy variable to measure whether a member signed a certain petition. Table A.1 shows the number of petitions in each Congress. It also shows the means and standard deviations of the petition variable by parties and by years. On average, a Congress member was less likely to sign a petition than to vote “yes” for a civil rights bill. Schickler explained that signing a discharge petition is costly for members of Congress, as it is a sign of violating “congressional norms” and intruding on “committee authority”. Only members who “cared enough” about a civil rights issue would sign a petition (p183). Hence, discharge petition signatures serve as a good measure of the *intensity* of a member’s preference.

Floor Speeches. We use a dummy variable to indicate whether a House member delivered at least one speech on the floor to support civil rights in a certain Congress. Floor speeches serve as an important way for House members to signal commitments to their constituents. Competition for speech time is fierce on the floor. Using precious time to deliver a speech to support civil rights indicates that a civil rights issue is one of the member’s top priorities. Table A.1 shows that the speech variable varies a lot over time. For instance, while in the 73rd Congress (1933-35), only about 2.5% members spoke about civil rights, in the 89th Congress (1965-67), about 22.5% Democrats and 16.1% Republicans delivered speeches supporting civil rights issues.

Bill Sponsorship. Bill Sponsorship is a count variable measuring how many civil rights bills a member initialized in a certain Congress. On average, one member sponsored less than 1 civil rights bill. The variation is large, as some members initialized many bills while most other members didn’t initialize any civil rights bills.

2.3 Limitations of the Original Analysis

In the Book, through analyzing the four types of data, Schickler finds that ever since the 1930s, more and more northern Democrats became advocates of civil rights. While at the beginning of the 1930s, northern Republicans were a bit more supportive of civil rights than northern Democrats, over the years, northern Democrats became more liberal while northern Republicans became more conservative. Since the 1940s, the *average* differences between the two parties’ positions became

³For a bill in a legislative committee, the waiting period is 20 days; for a bill in the Rules Committee, the waiting period is 7 days Schickler et al. (2010, p. 676).

larger and larger.⁴

Insightful as Schickler’s findings are, the average difference between the two parties’ positions reveals limited information on how old racial coalitions dissolved and new coalitions emerged. As “shifting partisan coalitions” played a central role in the transformation of racial politics, it is important to know how congressional members in different states and parties made coalitions with each other, and how their coalitions changed. For instance, in what states did congressional Democrats first join the new racial coalition? Democrats of what states were late joiners? Within a state, when did congressional Democrats and Republicans act together? When did they start to take opposite positions?

Also, the average racial position of a party says little about the distribution of racial positions within the party. Schickler (2016, p. 187) admits that “these findings still leave open the question of *which* northern Democrats and Republicans were particularly likely to back civil rights initiatives.” As racial realignment is a gradual procedure that initialized by locally oriented rank-and-file members, it is necessary to understand how members in individual states changed their positions over the years.

The proposed methodology provides an analytical tool to solve the questions listed above. We regard House members of a party in one state as the unit of analysis. By categorizing the party-state units into different groups and analyzing how the group membership changes dynamically over time, we can trace the gradual shift of racial coalitions as well as the gradual changes of party positions of individual states.

3 Model and Inference

3.1 The Model

The proposed dynamic DP model is a nonparametric and dynamic Bayesian clustering model. As is the DPMM, the dynamic DP model is a model for the latent group membership of observations and nonparametric in the sense that the number of latent clusters is not specified *a priori*. However, while the DPMM clusters cross-sectional units, the dynamic DP model is a model in which units move across groups over time. It is a dynamic model since the group membership of each unit changes over time, and the temporal shift of the membership is modeled as a Markov process. In each time period, a unit remains in the same group as the previous period or moves to a different group. If the unit moves, the group to which it is switching is determined by the Dirichlet process conditional on the group memberships in the previous period. Therefore, the proposed model is considered as a dynamic extension of DP.

To formally describe the model, let $i \in \{1, \dots, N\}$ denote a unit and $t \in \{1, \dots, T\}$ denote a time period in a panel data set. Also, let Y_{it} be observed (possibly multivariate) measurement

⁴The major findings of analyzing congressional data is in Chapter 8. Schickler also shows the results of analyzing the same four types of data in a earlier paper coauthored with Kathryn Pearson and Brian D. Feinstein (Schickler et al., 2010).

for unit i in time t . It is assumed that unit i in time t belongs to a latent group $g[it]$, where $g[it] \in \{1, 2, \dots\}$ denotes the latent group index of unit i in time t , and that Y_{it} is generated from a probabilistic model f with parameter $\Theta_{g[it]}$:

$$Y_{it} \sim f(\Theta_{g[it]}). \quad (1)$$

That is, Y_{it} and $Y_{i't'}$ share a common DGP if $g[it] = g[i't']$, but they may follow different DGPs otherwise. An appropriate prior distributions are placed on the model parameters Θ . A simple example of f is the Gaussian regression model with group-specific parameters:

$$Y_{it} \stackrel{\text{indep.}}{\sim} \mathcal{N}(X_{it}^\top \boldsymbol{\beta}_{g[it]}, \sigma_{g[it]}^2)$$

In this example, $\Theta_g = (\boldsymbol{\beta}_g, \sigma_g)$ and the parameters vary across latent groups. Thus, this regression model accounts for unobserved heterogeneity of the data generating process.

The dynamic DP models the generative process of group assignment, $g[it]$, in the generic data model defined in equation (1). First, the group membership in the initial period ($t = 1$) is generated by DP. Using the stick-breaking construction (Sethuraman, 1994), the generative process of $g[i1], i = 1, \dots, N$, is hierarchically defined as:

$$g[i1] \stackrel{\text{i.i.d.}}{\sim} \text{Discrete}(\{q_k\}_{k=1}^\infty) \quad (2)$$

$$q_k = \pi_k \prod_{l=1}^{k-1} (1 - \pi_l) \quad (3)$$

$$\pi_k \stackrel{\text{i.i.d.}}{\sim} \text{Beta}(1, \gamma) \quad (4)$$

where γ is the concentration parameter of the DP.

The key innovation of the proposed dynamic DP model is the dynamic generative process of latent groups for periods $t = 2, \dots, T$. Specifically, $g[it]$ follows a Markov process conditional on $(g[1, t-1], \dots, g[N, t-1])$. On one hand, similar to the sticky HDP-HMM model (Fox et al., 2011), unit i in time t stays in the same group as it was in time $t-1$ with probability p . Formally,

$$\Pr(g[it] = g[i, t-1] \mid g[i, t-1]) = p \quad (5)$$

where

$$p \sim \text{Beta}(\alpha_p, \beta_p). \quad (6)$$

On the other hand, with probability $1-p$, unit i 's group in time t follows the Dirichlet process,

but conditional on the cluster assignments in $t - 1$. That is,

$$g[it] \stackrel{\text{i.i.d.}}{\sim} \text{Discrete}(\{q_k^t\}_{k=1}^\infty) \quad (7)$$

$$q_k^t = \pi_k^t \prod_{l=1}^{k-1} (1 - \pi_l^t) \quad (8)$$

$$\pi_k^t \sim \text{Beta}(1 + n_k^{t-1}, \gamma + N - \sum_{l=1}^k n_l^{t-1}) \quad (9)$$

where n_k^{t-1} is the number of units in latent cluster k in period $t - 1$.

3.2 Intergenerational Chinese Restaurant Metaphor

The generative process for the latent groups is intuitively illustrated by the Chinese restaurant representation (Blackwell and MacQueen, 1973) of the Dirichlet process. The Dirichlet process is known to be equivalent to two constructive representation: The stick-breaking process and the Chinese restaurant process. The model described above is built on the former, because it is mathematically simpler. However, the Chinese restaurant representation provides far more intuitive illustration of the identical generative process. The name of the proposed model, the intergenerational Chinese restaurant process, comes from this representation.

Figure 1 illustrates the Chinese restaurant representation of the generative process in $t = 1$. In this representation, units are assigned to latent groups sequentially. At the beginning, unit 1 is (arbitrarily) assigned to group 1. For each subsequent unit, the probability that the unit assigned to a group is proportional to the number of the units already assigned to that group, and the probability that a unit creates a new cluster is proportional to prior parameter γ . For example, unit 2 goes to group 1 with probability $1/(1 + \gamma)$ or creates a new cluster (cluster 2) with probability $\gamma/(1 + \gamma)$ (the top panel of Figure 1). If, unit 2 is assigned to group 2 in the realized state, the next unit (unit 3) is assigned to group 1 or 2 with probability $1/(2 + \gamma)$, or forms a new group (group 3) with probability $\gamma/(2 + \gamma)$ (the middle panel of Figure 1). Furthermore, given that unit 1 is assigned to group 1 and units 2 and 3 are assigned to group 2, unit 4 is assigned to group 1 with probability $1/(3 + \gamma)$, group 2 with probability $2/(3 + \gamma)$, and group 3 with probability $\gamma/(3 + \gamma)$, respectively (the bottom panel of Figure 1). This process is equivalent to the model defined by equations (2) through (4) for latent groups in $t = 1$.

The dynamic process in the proposed model defined by equations (5) through (9) has analogous interpretation. Figure 2 illustrates the generative process of latent groups in $t = 2$. This process is the Chinese restaurant process given the group assignments in $t = 1$ with some stickiness added.

The top panel of Figure 2 illustrates the group assignment of the first unit in period 2. Here, only for the purpose of illustration, it is assumed that two units are in group 1 and the other two units are in group 2 in period 1. In period 2, unit 1 has some stickiness to group 1, because it is in group 1 in the previous period. This stickiness is represented by probability parameter p . However,

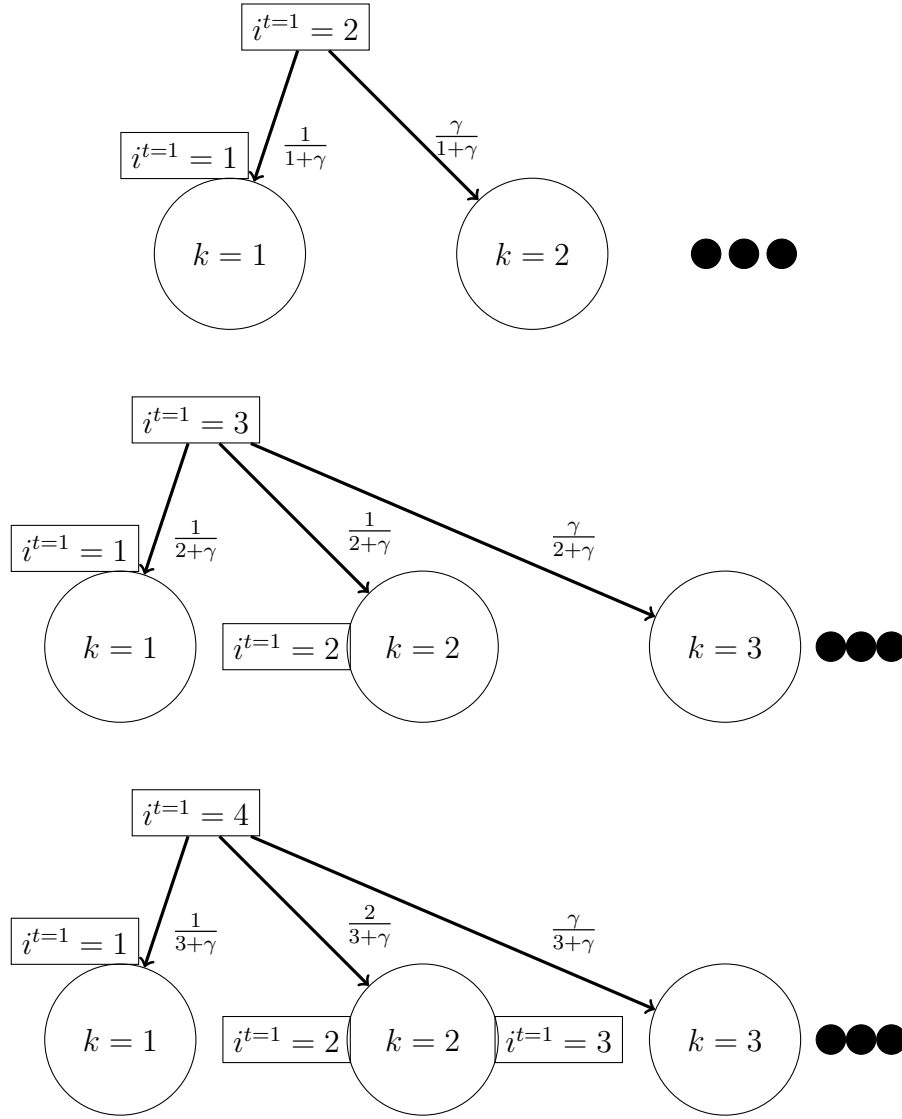


Figure 1: Chinese Restaurant Process. This figure illustrates the generative process of the latent cluster membership in $t = 1$. The fractions near arrows are the probabilities of cluster assignment. The top, the middle and the bottom panel show the probability that unit 2, 3, and 4 are assigned to each cluster, respectively. The probability that a unit assigned to a cluster is proportional to the number of previously assigned units, and the probability that a unit creates a new cluster is proportional to parameter γ .

to model possible change in group assignment for unit 1, unit 1 enters the Chinese restaurant process given group assignments in period 1 with probability $1 - p$. This Chinese restaurant process works as follows: the probability of unit 1 being assigned to group k is proportional to the number of the units assigned to group k in period 1. For example, the probability that unit 1 will be assigned to group 2 is $(1 - p) \times 2/(4 + \gamma)$, because unit 1 enters the Chinese restaurant process with probability $1 - p$ and if it does, it will be assigned to group 2 with probability proportional to 2, which is the number of the units assigned to group 2 in period 1. On the other hand, unit 1 may still stay in group 1 even if it enters the Chinese restaurant process. The probability of that happening is $(1 - p) \times 1/(4 + \gamma)$, because it follows the Chinese restaurant process with probability

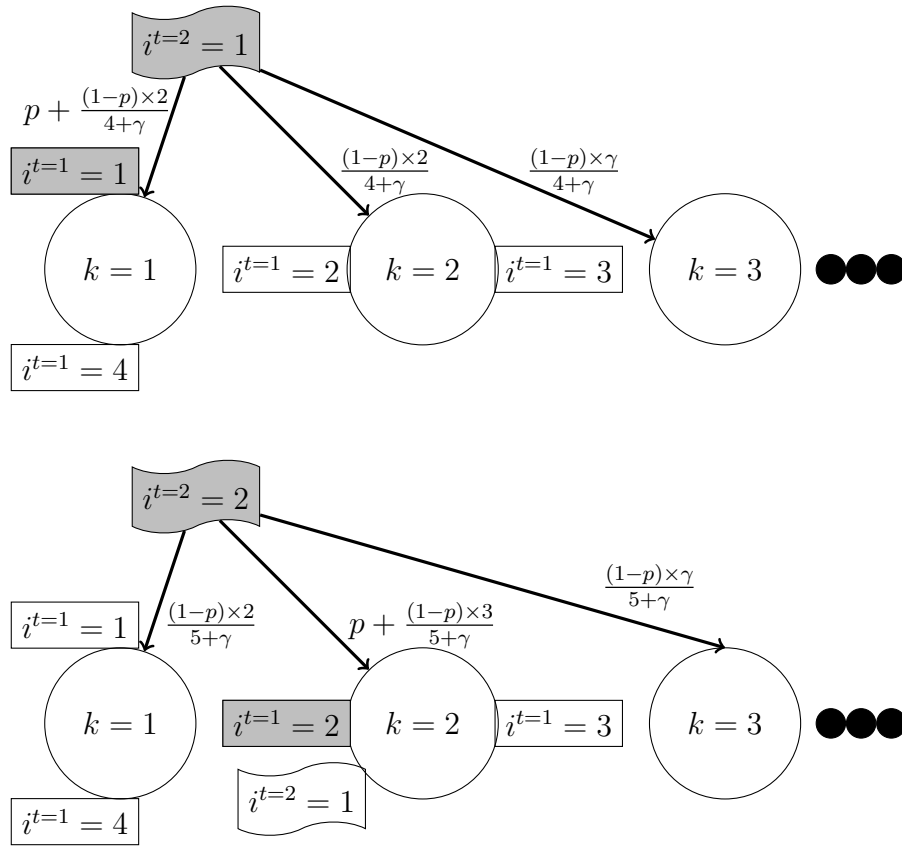


Figure 2: Intergenerational Chinese Restaurant Process. This figure illustrates the generative process of the latent cluster membership in $t = 2$. The fractions near arrows are the probabilities of cluster assignment. The top and bottom panels present the probability that unit 1 and 2 in period 2 are assigned to each cluster, respectively. The probability that a unit assigned to a cluster depends on cluster assignments in $t = 1$ and units already assigned to clusters in $t = 2$. The probability is proportional to $1 - p$ times the number of units in $t = 1$ and 2, and the probability that a unit creates a new cluster is proportional to $1 - p$ times parameter γ . In addition, there is stickiness that stickiness parameter p is added to the probability that a unit is assigned to the cluster it occupied in $t = 1$.

$1 - p$ and the conditional probability of group 1 is proportional to 2, which is the number of the units in group 1 in period 1. Since unit 1 is sticky to group 1 with probability p , the total probability of unit 1 being assigned to group 1 is $p + (1 - p)2/(4 + \gamma)$.

The group assignment of the other units again follows the Chinese restaurant process conditional on the period 1 group assignment, with stickiness parameter p . The bottom panel of Figure 2 illustrates the group assignment of unit 2 in period 2, assuming that unit 1's assignment in period 2 is group 2. Again, unit 2 is sticky to the cluster to which the same unit was assigned in period 1. Therefore, the probability that unit 2 goes to cluster 2 is p plus the probability determined by the Chinese restaurant process. The difference from the case of unit 1 is that the Chinese restaurant process for unit 2 involves unit 1 in period 2. That is, the probability that unit 2 is assigned to cluster 2 should be $p + (1 - p) \times 3/(5 + \gamma)$, because two units in period 1 and unit 1 in period 2 are assigned to cluster 2.

The probability of each of the other groups is simply determined by the number of units in that group. Since group 1 has two units from period 1, the conditional probability that unit 2 belongs to group 1 is $2/(5 + \gamma)$. On the other hand, the conditional probability of a new group formation (group 3 in the figure) is $\gamma/(5 + \gamma)$. The marginal probabilities are obtained by multiplying the conditional probability by $1 - p$, the probability that the unit enters the Chinese restaurant process.

3.3 Related Models

Among other dynamic DP models, the proposed dynamic DP model is most closely related to Caron et al. (2012). However, equation (9) is different in that cluster assignment in t depends on the assignments of all units in $t - 1$, instead of the units whose assignments are not deleted in t . Hence, our model makes stronger temporal dependence between generations, because in case of $p = 0$ where there is no stickiness, the group assignment in t is still influenced by the assignment in $t - 1$ in our model whereas it is not in Caron et al. (2012). In the CRP metaphor, units in t in our model enter the restaurant with tables occupied by units in $t - 1$, but in Caron et al. (2012), they enter an empty restaurant.

Iorio et al. (2023) develops another dynamic DP, in which π 's in Equation (8) follow an autoregressive process. In their model, the cluster assignment does not depend on the assignment in the previous period. In addition, the cluster specific parameters are not shared across time. In Huang et al. (2015), units stay in the same cluster for the entire life. They join a cluster and leave from the data set after a certain length of time.

The time-sensitive Dirichlet process mixture model developed by Zhu et al. (2005) does not have a stickiness parameter. Dunson (2006); Ren et al. (2010); Zhang et al. (2010); Das et al. (2021) are also a dynamic DP without stickiness. Their dynamics is as follows: DP in period t is a mixture of the DP in $t - 1$ and an entirely new DP (units in t go into the same CRP as $t - 1$ or a brand-new CRP).

3.4 Markov Chain Monte Carlo Algorithm for Estimation

In this part, we introduce the Markov Chain Monte Carlo (MCMC) algorithm to estimate the model constructed through the dynamic stick-breaking process. In general, we use the blocked Gibbs sampling algorithm. It includes two major steps. One step is to sample the group-specific parameter Θ_g that defines the data generating function $f(\Theta_k)$ for $k = 1, 2, \dots$, conditioned on the current group assignment $g[it] = k$. Sampling Θ_k from its posterior distribution is simple when the prior distribution for Θ_k is conjugate. The second step is to sample group assignments $g[it]$ for $i = 1, 2, \dots, N$ and $t = 1, 2, \dots, T$ conditioned on Θ_k for $k = 1, 2, \dots$. Sampling $g[it]$ is relatively complicated, as the posterior distribution is conditioned on both cross-sectional and cross-time relationships with other units. We combine the forward-backward algorithm designed for change-point models (Chib, 1998) and the truncation approximation approach developed for stick-breaking priors (Ishwaran and James, 2001) to sample $g[it]$. In what follows, we describe the details for this MCMC algorithm.

Following the truncation approximation approach of Ishwaran and James (2001), at the beginning of the algorithm, we set an arbitrarily large number K at which we truncate the number of groups. In the posterior distributions of $g[it]$ for all $i = 1, 2, \dots, N$ and $t = 1, 2, \dots, T$, the probabilities for most groups will be zero. In this way, the model estimates the number of groups automatically.

With K set, we initialize the starting values of $g[it]$ for all $i = 1, \dots, N$ and $t = 1, \dots, T$. Then, each iteration of the Gibbs sampler proceeds as follows.

First, we update Θ_k for $k = 1, 2, \dots, K$. The posterior distribution of Θ_k is only conditioned on X_{it} and Y_{it} for units in group k . The sampling algorithm for Θ_k is specific to the form of function $f(\Theta_{g[it]})$. If possible, set the prior for Θ_k a conjugate prior to make the sampling simple.

Second, we sample p , the probability that a unit stays in the same group as before and does not go through a new round of group assigning process. To make the sampling of p simpler, we first introduce a series of intervening dummy variables d_{it} for $i = 1, 2, \dots, N$ and $t = 1, 2, 3, \dots, T$ to indicate whether a specific unit directly stays in the same group as before. The posterior distribution of d_{it} is only conditioned on $g[it]$, $g[i, t - 1]$, and p . When $g[it] \neq g[i, t - 1]$, d_{it} must be 0; when $g[it] = g[i, t - 1]$, unit i may directly stay in the same group as before, or may be assigned to the same group through the new round of group assigning process. Specifically,

$$p(d_{it} = 1) = \begin{cases} 0 & \text{if } g[it] \neq g[i, t - 1]; \\ \frac{p}{p + (1-p)q_k^t} & \text{if } g[it] = g[i, t - 1] = k. \end{cases}$$

As we set the prior distribution for p a Beta distribution, conditioned on d_{it} for all $i = 1, 2, \dots, N$ and $t = 1, 2, \dots, T$, the posterior distribution of p is also a Beta distribution. Specifically, we sample p as follows:

$$p \sim \text{Beta}(\alpha_p + N_1, \beta_p + N_2)$$

$$N_1 = \sum_{i=1}^N \sum_{t=2}^T d_{it}$$

$$N_2 = \sum_{i=1}^N \sum_{t=2}^T (1 - d_{it})$$

Third, we update the stick-breaking weight π_k^t and q_k^t for $k = 1, 2, \dots, K$ and $t = 1, 2, \dots, T$. As we set K to be an arbitrarily large number to approximate the stick-breaking prior with infinite number of groups, the posterior distribution of π_k^t simply becomes a Beta distribution:

$$\pi_k^t \sim \text{Beta}(1 + n_k^{t-1} + n_k^t, \gamma + \sum_{l=k+1}^K n_l^{t-1} + \sum_{l=k+1}^K n_l^t)$$

$$n_k^{t-1} = \sum_{i=1}^N \mathbb{I}(g[i, t - 1] = k)$$

$$n_k^t = \sum_{i=1}^N (1 - d_{it}) \mathbb{I}(g[it] = k)$$

Once π_k^t is updated, calculate q_k^t :

$$q_k^t = \pi_k^t \prod_{l=1}^{k-1} (1 - \pi_l^t)$$

The last step is to sample group assignments $g[it]$ for $i = 1, 2, \dots, N$ and $t = 1, 2, \dots, T$. Let $\mathbf{g}[t] \equiv (g[1t], g[2t], \dots, g[Nt])'$, $\mathbf{q}^t \equiv (q_1^t, q_2^t, \dots, q_K^t)'$, $\mathbf{Y}_t \equiv (Y_{1t}, Y_{2t}, \dots, Y_{Nt})'$. Following Chib (1998), we sample $\mathbf{g}[T], \mathbf{g}[T-2], \dots, \mathbf{g}[1]$ in turn:

$$Pr(g[it] = k | \mathbf{g}[T], \dots, \mathbf{g}[t+1], p, \mathbf{q}^T, \dots, \mathbf{q}^1, \mathbf{Y}_T, \dots, \mathbf{Y}_1, \Theta_1, \dots, \Theta_K) \quad (10)$$

$$\propto \underbrace{Pr(g[i, t+1] | p, g[it] = k, \mathbf{q}^{t+1})}_{\text{part 1}} \underbrace{Pr(g[it] = k | p, \mathbf{q}^t, \dots, \mathbf{q}^1, \mathbf{Y}_t, \dots, \mathbf{Y}_1, \Theta_1, \dots, \Theta_K)}_{\text{part 2}} \quad (11)$$

The equation above shows a decomposition of the conditional posterior distribution of $g[it]$ given the observed data, the model parameters, and group assignments for period $t+1$ through T . Part 1 is the probability of $g[i, t+1]$ given that $g[i, t]$ is in group k . Then part 1 is :

$$Pr(g[i, t+1] = l | p, g[it] = k, q_l^{t+1}) = (1-p)q_l^{t+1} + p\mathbb{I}(l = k)$$

Part 2 is the conditional probability of $g[it] = k$ given the model parameters and the data from period 1 up to period t . Part 2 can be further decomposed as:

$$Pr(g[it] = k | p, \mathbf{q}^t, \dots, \mathbf{q}^1, \mathbf{Y}_t, \dots, \mathbf{Y}_1, \Theta_1, \dots, \Theta_K) \\ \propto \underbrace{f(Y_{it} | g[it] = k, \mathbf{Y}_{t-1}, \dots, \mathbf{Y}_1, \Theta_1, \dots, \Theta_K)}_{\text{part a}} \underbrace{Pr(g[it] = k | p, \mathbf{q}^t, \dots, \mathbf{q}^1, \mathbf{Y}_{t-1}, \dots, \mathbf{Y}_1, \Theta_1, \dots, \Theta_K)}_{\text{part b}}$$

Part a is the distribution of outcome variable Y_{it} conditioned on contemporary group assignment $g[it]$, outcome variables of former periods, and group-specific parameters Θ_k for $k = 1, 2, \dots, K$. Given the group assignment $g[it] = k$ and parameter for the group Θ_k , the distribution of Θ_k is conditionally independent of the data in the other periods:

$$f(Y_{it} | g[it] = k, \mathbf{Y}_{t-1}, \dots, \mathbf{Y}_1, \Theta_1, \dots, \Theta_K) = f(Y_{it} | \Theta_k)$$

Compared to the conditional distribution of $g[it]$ in Part 2, in Part b $g[it]$ is no longer conditioned on \mathbf{Y}_t . Part b can be further decomposed as:

$$Pr(g[it] = k | p, \mathbf{q}^t, \dots, \mathbf{q}^1, \mathbf{Y}_{t-1}, \dots, \mathbf{Y}_1, \Theta_1, \dots, \Theta_K) \\ = \sum_l \underbrace{Pr(g[it] = k | p, q_k^t, g[i, t-1] = l)}_{\text{part c}} \underbrace{Pr(g[i, t-1] = l | p, \mathbf{q}^{t-1}, \dots, \mathbf{q}^1, \mathbf{Y}_{t-1}, \dots, \mathbf{Y}_1, \Theta_1, \dots, \Theta_K)}_{\text{part d}}$$

Now one can immediately observe the recursive structure between the equation above and part 1 and part 2 in equation (11). Therefore, the standard forward recursion algorithm computes the conditional posterior given by equation (10), and the backward sampling algorithm generates MCMC draws from the conditional posterior of the group assignment.

4 Empirical Application

In this section, we apply the proposed methodology to study shifting party coalitions on civil rights. First, we introduce the four measures of civil rights positions Schickler (2016) uses. Then, we explain how to apply the dynamic DP model to the empirical example. We introduce the specific statistical model that we use to analyze the data of civil rights positions. After this, we show the results of statistical analysis. Besides confirming Schickler’s original argument that racial realignment was a gradual procedure that started ever since the New Deal period, the new results provide more information on how old coalitions dissolved and new coalitions emerged. The new results also provide inference on how positions of congressional members in individual states changed over the years.

4.1 Statistical Model for Analyzing Racial Realignment

Here we introduce the specific statistical model we use to analyze the data of congressional action. We regard a party’s House members in one state as the basic unit of analysis, assuming members from the same state and in the same party share the same constraint imposed by their constituencies. We use the IgCRP model along with the voting block model of Spirling and Quinn (2010) as the data model, to group party-state units and to model the evolvement of group memberships over the years. Party-state units in the same group share the group-specific parameters that we use to model the four types of legislative behaviors. In this way, we analyze the four types of data measuring House members’ civil rights positions in a unified framework.

Specifically, let $t = 73, 74, \dots, 92$ index the 73rd Congress (1933-35) to the 92nd Congress (1971-73), $s = 1, 2, \dots, 50$ index 50 states, and D and R represent the Democratic Party and the Republican Party respectively. Let $i = 1, 2, \dots, I_{st}^D$ (or I_{st}^R) represent a Democratic/Republican House member from state s in the t th Congress, and $j = 1, 2, \dots, J_t^V / J_t^P$ index a civil rights vote or petition in the t th Congress.

Let $g^D[st]$ represent the group the Democratic Party of state s in the t th Congress belongs to. Similarly, $g^R[st]$ represents the group of the Republican party. We assume $g^D[st]$ and $g^R[st]$ share the same IgCRP prior:

$$g^D, g^R \sim \text{IgCRP}(\gamma, \alpha_p, \beta_p)$$

Roll-call Votes. Let V_{istj}^D represent the outcome of roll-call vote j in the t th Congress for Democratic member i from state s . Similarly, V_{istj}^R represents the roll-call vote of a Republican

member. We assume V_{istj}^D and V_{istj}^R follow the following distributions:

$$V_{istj}^D \stackrel{ind.}{\sim} \text{Bernoulli}(\theta_{g^D[st]})$$

$$V_{istj}^R \stackrel{ind.}{\sim} \text{Bernoulli}(\theta_{g^R[st]})$$

For $g^D[st] = 1, 2, \dots, k \dots$ and $g^R[st] = 1, 2, \dots, k \dots$,

$$\theta_k \sim \text{Beta}(\alpha_\theta, \beta_\theta)$$

Discharge Petition Signatures. Let P_{istj}^D and P_{istj}^R represent whether member i from state s signed petition j in the t th Congress. D and R index Democrats and Republicans. We assume P_{istj}^D and P_{istj}^R follow the following distributions:

$$P_{istj}^D \stackrel{ind.}{\sim} \text{Bernoulli}(\eta_{g^D[st]})$$

$$P_{istj}^R \stackrel{ind.}{\sim} \text{Bernoulli}(\eta_{g^R[st]})$$

For $g^D[st] = 1, 2, \dots, k \dots$ and $g^R[st] = 1, 2, \dots, k \dots$,

$$\eta_k \sim \text{Beta}(\alpha_\eta, \beta_\eta)$$

Floor Speeches. Let S_{ist}^D and S_{ist}^R represent whether Democratic or Republican member i from state s delivered at least one floor speech to support civil rights. Their distributions are:

$$S_{ist}^D \stackrel{ind.}{\sim} \text{Bernoulli}(\omega_{g^D[st]})$$

$$S_{ist}^R \stackrel{ind.}{\sim} \text{Bernoulli}(\omega_{g^R[st]})$$

For $g^D[st] = 1, 2, \dots, k \dots$ and $g^R[st] = 1, 2, \dots, k \dots$,

$$\omega_k \sim \text{Beta}(\alpha_\omega, \beta_\omega)$$

Bill Sponsorship. Finally, let B_{ist}^D and B_{ist}^R represent the number of civil rights bills member i from state s sponsored in the t th Congress. B_{ist}^D and B_{ist}^R follow the following distributions:

$$B_{ist}^D \stackrel{ind.}{\sim} \text{Poisson}(\lambda_{g^D[st]})$$

$$B_{ist}^R \stackrel{ind.}{\sim} \text{Poisson}(\lambda_{g^R[st]})$$

For $g^D[st] = 1, 2, \dots, k \dots$ and $g^R[st] = 1, 2, \dots, k \dots$,

$$\lambda_k \sim \text{Gamma}(a_\lambda, b_\lambda)$$

4.2 MCMC Algorithm

As introduced in Section 3.4, the MCMC algorithm includes two major steps. One step is to sample the group-specific parameters, here θ_k , η_k , ω_k , and λ_k , for $k = 1, 2, \dots$, conditioned on group assignments $g^D[st]$ and $g^R[st]$ for $s = 1, 2, \dots, 50$ and $t = 73, 74, \dots, 92$. We set priors for the group-specific parameters conjugate priors. Conditioned on group assignments, posterior distributions of θ_k , η_k , ω_k are Beta distributions, and the posterior distribution of λ_k is a Gamma distribution. The second step is to sample group assignments $g^D[st]$ and $g^R[st]$ for $s = 1, 2, \dots, 50$ and $t = 73, 74, \dots, 92$, conditioned on the group-specific parameters. We combine the forward-backward approach for change-point models and the truncation approximation approach for stick-breaking priors to sample group assignments. Details of the algorithm are shown in SI B.

4.3 Empirical Results

In this part, we first check whether our results are consistent with the original findings of Schickler (2016). Beyond this, we present the changes of the two parties' positions in individual states, cross-party coalition within a state, as well as cross-state coalition within a party.

Differences Across Parties. We use posterior means of outcome variables, $\{\theta_{g^D[st]}, \theta_{g^R[st]}\}$, $\{\eta_{g^D[st]}, \eta_{g^R[st]}\}$, $\{\omega_{g^D[st]}, \omega_{g^R[st]}\}$, $\{\lambda_{g^D[st]}, \lambda_{g^R[st]}\}$ to measure a party's position in state s and the t th Congress. The posterior mean carries easily-interpreted empirical meaning. For instance, $\theta_{g^D[st]}$ means the probability that in the t th Congress a Democratic member from state s would vote "yes" for a civil rights roll call; $\lambda_{g^D[st]}$ means the average number of civil rights bill Democratic members in state s sponsored in the t th Congress.

Then we investigate how the difference between the two parties' positions changed over time. To compare our results with the original findings, We take the average of northern states and southern states separately to check whether the two parties' positions had started to diverge longer before the 1960s in the North.

We present the results in Figure 3. Consistent with the original findings, we find that ever since the 1940s, on average, northern Democrats were more supportive of civil rights than northern Republicans, and the gap of the two parties in the North kept increasing until the 1960s. Blue solid lines in each panel take positive values on the y-axis, which indicates that northern Democrats are more supportive of civil rights than Republicans from the same states. While the Democratic Party was only moderately different from the Republican Party in terms of roll call votes, for all the other three measures, the gaps were substantially large.

The results also automatically capture the well-known situation in the South that southern Democrats stood firmly against civil rights legislation. In the 1930s, the average differences between Democrats and Republicans in the South were negative. However, since the 1940s, the partisan gap decreased gradually, and by the end of the 1950s, the differences became near 0. This change was partly driven by the shift of southern Republicans' positions and partly driven by the gradually dissolved Democratic coalition in the South. We will further explore this change

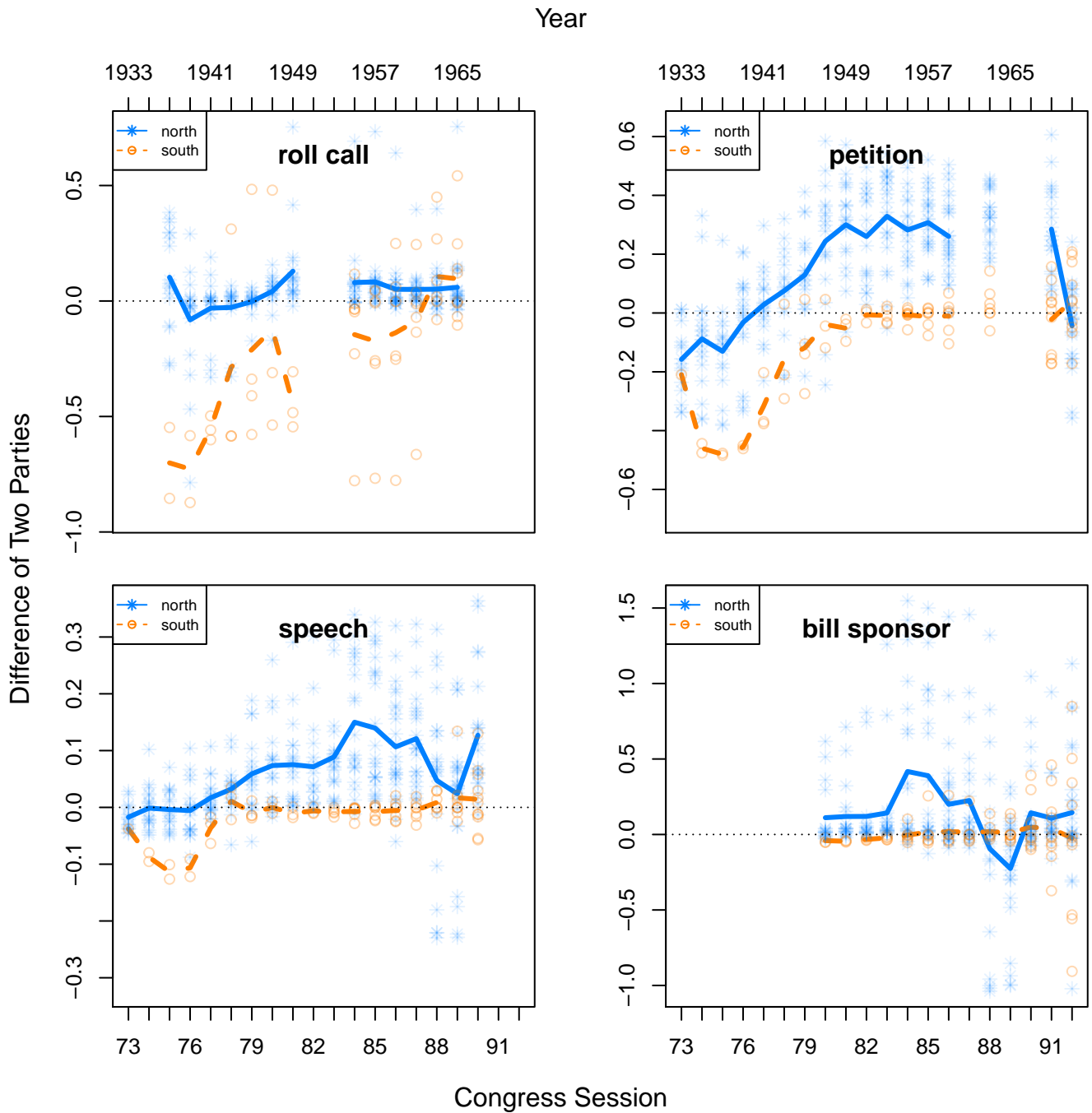


Figure 3: Differences between the Democratic Party and the Republican Party on Civil Rights Position in Each Congress. A positive value in the y-axis indicates that Democrats are more supportive of civil rights. Each point represents the posterior mean of an outcome variable in the corresponding state and Congress. The solid and dotted lines represent averages of northern and southern states respectively.

in the following analysis.

Beyond the average differences between Democrats and Republicans, our model also reveals how the partisan differences in individual states evolved over time. Figure 3 shows that even viewed from the perspective of roll-call votes, a measure only moderately captures the two parties' difference, the pattern of partisan differences started to change in northern states. In the 1930s,

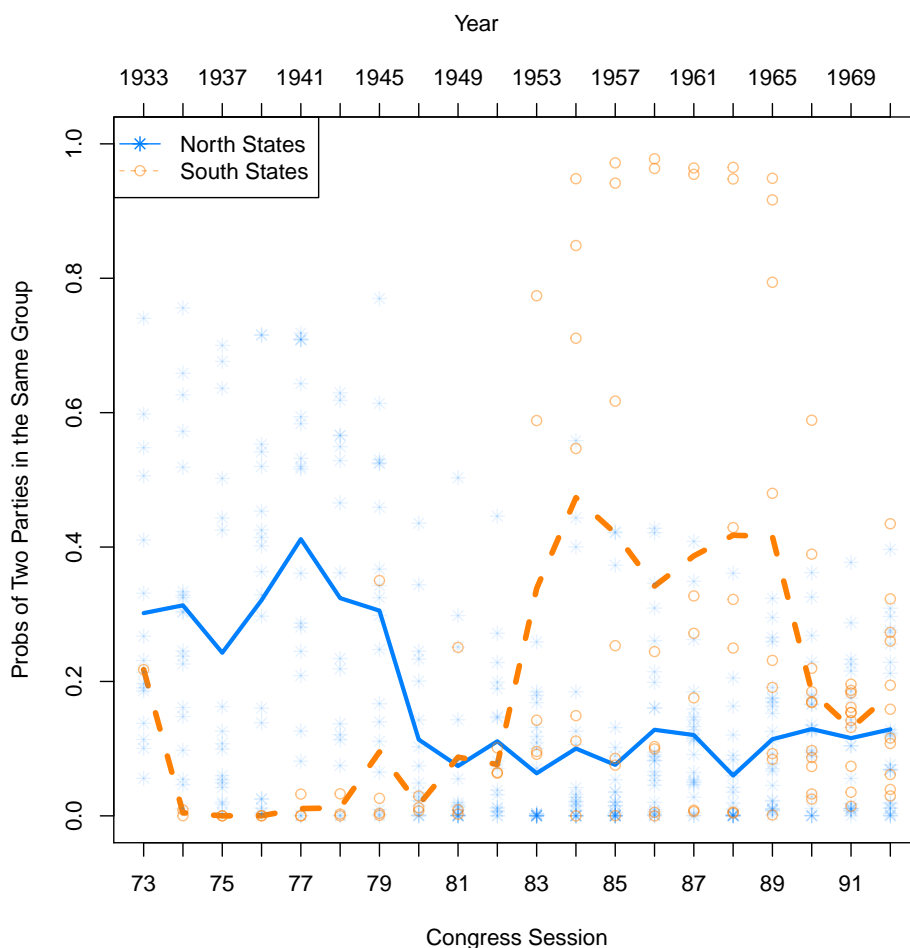


Figure 4: Probabilities that the Democratic Party and the Republican Party Are in the Same Group in Supporting Civil Rights in Each Congress. Each point represents the probability of an individual state. The solid and dotted lines represent averages of northern and southern states respectively.

there were many northern states where Republicans were more supportive of civil rights than Democrats. Representative states were Missouri, Nebraska, and Oregon.⁵ In the 1940s, the number of such states decreased gradually. By the end of the 1940s, in few states the partisan differences were negative. The proposed methodology also helps identify states where Democrats pioneered in advancing civil rights. New York, California, and Illinois are such states.

Cross-party Coalitions within a State. Besides posterior means of outcome variables, we also use the probability that the Democratic Party and the Republican Party are in the same group to measure how the two parties differed in each state and each Congress. In addition, we calculate the average probabilities of northern and southern states separately.

As shown in Figure 4, the average probability of northern states declines fast in the 1940s. Before, in some northern states such as Connecticut and New Jersey, the probabilities are as high as 0.7 to 0.8, meaning the Democrats and Republicans took very similar positions on civil rights issues. Since the 1950s, in nearly all northern states the probabilities are below 0.5. In many

⁵Note that the “North” is defined all the non-South states. See footnote 1

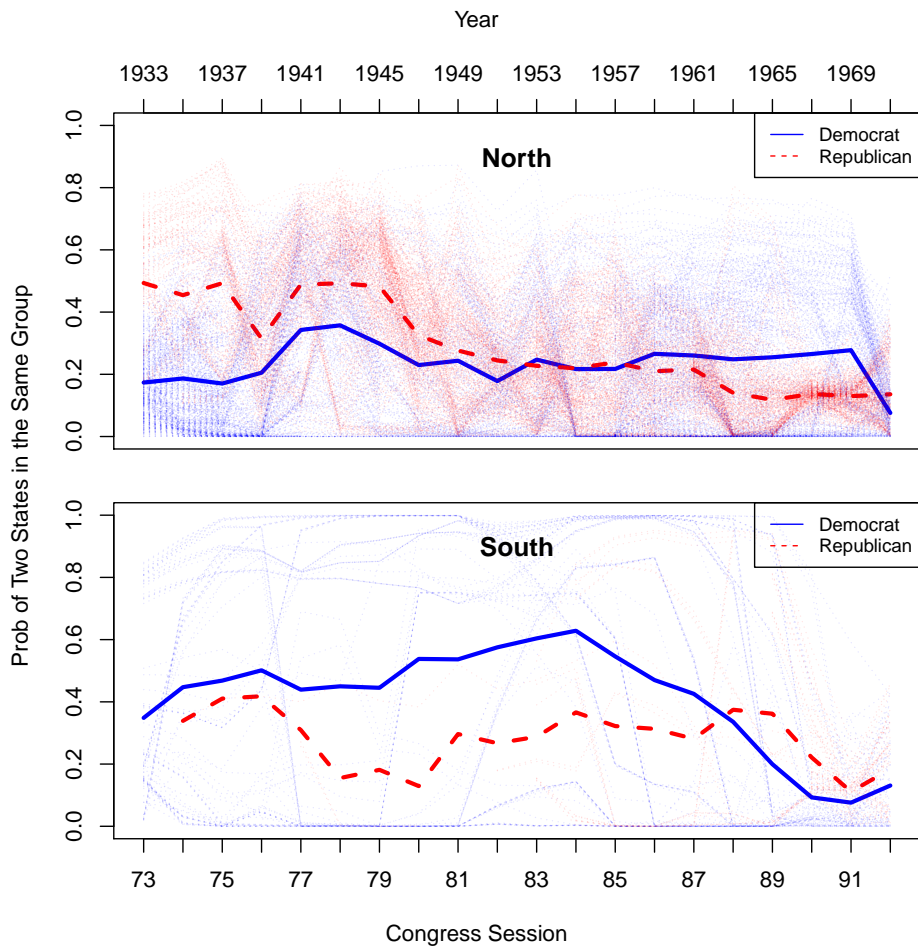


Figure 5: Probabilities that a Party in Two Different States are in the Same Groups in Each Congress. Light lines represent the probabilities of two specific states. Thick lines represent average probabilities of the North and the South.

states, the probabilities are near zero, meaning Democrats and Republicans took very different positions on civil rights issues.

In the South, the Democratic Party and the Republican Party used to hold sharply different attitudes toward civil rights. The probabilities of southern states are near zero in most Congresses before 1950. In the 1950s, the probabilities increase fast. During some Congresses of the 1950s and the 1960s, in states such as Virginia and North Carolina, the probabilities approach 1, meaning that in these states southern Democrats and Republicans hardly disagreed with each other on civil rights issues.

Cross-state Coalitions within a Party. Besides coalitions across parties, we also illustrate within-party coalitions across different states and how the coalitions changed over time. In Figure 5, we show the probabilities that Democrats/Republicans in different states are in the same group. We also show the average probabilities of northern states and southern states separately. The probabilities measure the degree of party solidarity.

In the 1930s, northern Republicans were more unified than northern Democrats. Starting from the end of the 1930s, northern Democrats became more likely to act together. While Democratic

solidarity decreased slightly in the first half part of the 1940s, after the 80th Congress (1947-49), the degree of solidarity restarted to increase slowly again. At the same time, Republican solidarity decreased gradually over time. By the end of the 1950s, northern Republicans had become less united than northern Democrats.

In contrast, in the South, Democratic used to be firmly unified. As northern states started to act together to support civil rights, southern Democrats became more united in opposing civil rights legislation. However, since the 1950s, the southern solidarity of the Democratic Party started to dissolve gradually. By the end of the 1960s, the average probability that Democrats in two southern states are in the same group has declined to around 0.1.

5 Concluding Remarks

Most societal changes are evolutionary rather than revolutionary, that is, gradual rather than discontinuous. As our empirical analysis indicates, the Democratic Party's shift to the pro-civil rights position began in the 1930s and continued gradually until the 1960s. Before the landmark presidential election of 1964, the Southern Democrats' unity against civil rights started eroding in the 1950s. To properly analyze evolutionary changes, one needs to model a continuing process of gradual shifts in a data generating process.

This paper presents a novel dynamic Dirichlet mixture model to achieve that goal. In the proposed model, units are clustered into latent groups, and the group assignment is assumed to be a Markov process. In each time period, a unit remains in the same group as the previous period with a probability controlled by a stickiness parameter, or moves to another group with a probability determined by the Chinese restaurant process conditional on the group assignments in the previous period. Due to this group assignment process, the proposed dynamic DP model can properly model a gradual structural change.

The proposed model is widely applicable to any data set with the repeated measurement of multiple units, and thus its use would benefit many social scientists who are interested in temporal dynamics of unobserved heterogeneity. To facilitate this, we are developing free, open-source, and easy-to-use software for implementing the MCMC algorithm. We intend to create a package in the R system consisting of pre-written functions for estimation algorithms of the model so that applied quantitative social scientists can simply use the package to conduct Bayesian analysis using the model.

References

- Antoniak, C. E. (1974), “Mixtures of Dirichlet Processes with Applications to Bayesian Nonparametric Problems,” *The Annals of Statistics*, pp. 1152–1174.
- Blackwell, D., and MacQueen, J. B. (1973), “Ferguson distributions via Pólya urn schemes,” *The Annals of Statistics*, 1(2), 353–355.
- Califano, J. (1992), *The Triumph and Tragedy of Lyndon Johnson*, New York: Simon and Schuster.
- Caron, F., Davy, M., and Doucet, A. (2012), “Generalized Polya Urn for Time-varying Dirichlet Process Mixtures,” .
- Chen, A. S., Mickey, R. W., and Houweling, R. P. V. (2008), “Explaining the Contemporary Alignment of Race and Party: Evidence from California’s 1946 Ballot Initiative on Fair Employment,” *Studies in American Political Development*, 22, 204–28.
- Chib, S. (1998), “Estimation and Comparison of Multiple Change-point Models,” *Journal of Econometrics*, 86(2), 221–241.
- Das, K., Ghosh, P., and Daniels, M. J. (2021), “Modeling Multiple Time-Varying Related Groups: A Dynamic Hierarchical Bayesian Approach With an Application to the Health and Retirement Study,” *Journal of the American Statistical Association*, 116(534), 558–568.
- Dunson, D. B. (2006), “Bayesian dynamic modeling of latent trait distributions,” *Biostatistics*, 7(4), 551–568.
- Edsall, T., and Edsall, M. (1992), *Chain Reaction: The Impact of Race, Rights, and Taxes on American Politics*, New York: Norton.
- Farhang, S., and Katznelson, I. (n.d.), “The Southern Imposition: Congress and Labor in the New Deal and Fair Deal,” *Studies in American Political Development*, 19, 1–30.
- Ferguson, T. S. (1973), “A Bayesian Analysis of Some Nonparametric Problems,” *The Annals of Statistics*, 1(2), 209–230.
- Fox, E. B., Sudderth, E. B., Jordan, M. I., and Willsky, A. S. (2011), “A Sticky HDP-HMM with Application to Speaker Diarization,” *The Annals of Applied Statistics*, 5(2A), 1020–1056.
- Garmines, E. G., and Stimson, J. A. (1989), *Issue Evolution: Race and the Transformation of American Politics*, Princeton, NJ: Princeton University Press.
- Huang, R., Zhu, F., and Heng, P.-A. (2015), The dynamic chinese restaurant process via birth and death processes,, in *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 29.

- Imai, K., and Tingley, D. (2012), “A statistical method for empirical testing of competing theories,” *American Journal of Political Science*, 56(1), 218–236.
- Iorio, M. D., Favaro, S., Guglielmi, A., and Ye, L. (2023), “Bayesian nonparametric mixture modeling for temporal dynamics of gender stereotypes,” *The Annals of Applied Statistics*, 17(3), 2256 – 2278.
- Ishwaran, H., and James, L. F. (2001), “Gibbs Sampling Methods for Stick-Breaking Priors,” *Journal of the American Statistical Association*, 96(453), 161–172.
- Kim, I. S., Liao, S., and Imai, K. (2020), “Measuring Trade Profile with Granular Product-level Trade Data,” *American Journal of Political Science*, 64(1), 102–117.
- Kyung, M., Gill, J., and Casella, G. (2011), “New Findings from Terrorism Data: Dirichlet Process Random-Effects Models for Latent Groups,” *Journal of the Royal Statistical Society Series C: Applied Statistics*, 60(5), 701–721.
- McAdam, D., and Kloos, K. (2014), *Deeply Divided: Racial Politics and Social Movements in Post- War America*, New York: Oxford University Press.
- Pang, X., Friedman, B., Martin, A. D., and Quinn, K. M. (2012), “Endogenous Jurisprudential Regimes,” *Political Analysis*, 20(4), 417–436.
- Park, J. H. (2010), “Structural Change in U.S. Presidents’ Use of Force,” *American Journal of Political Science*, 54(3), 766–782.
- Ren, L., Dunson, D., Lindroth, S., and Carin, L. (2010), “Dynamic Nonparametric Bayesian Models for Analysis of Music,” *Journal of the American Statistical Association*, 105(490), 458–472.
- Schickler, E. (2016), *Racial Realignment: The Transformation of American Liberalism, 1932-1965*, Princeton, NJ: Princeton University Press.
- Schickler, E., Pearson, K., and Feinstein, B. D. (2010), “Congressional Parties and Civil Rights Politics from 1933 to 1972,” *The Journal of Politics*, 72(3), 672–689.
- Sethuraman, J. (1994), “A Constructive Definition of Dirichlet Priors,” *Statistica Sinica*, pp. 639–650.
- Spirling, A., and Quinn, K. (2010), “Identifying intraparty voting blocs in the UK House of Commons,” *Journal of the American Statistical Association*, 105(490), 447–457.
- Teh, Y. W. (2010), “Dirichlet Process,” in *Encyclopedia of Machine Learning*, New York: Springer, pp. 280–287.

Weir, M. (2005), “States, Race, and the Decline of New Deal Liberalism,” *Studies in American Political Development*, .

Zhang, J., Song, Y., Zhang, C., and Liu, S. (2010), Evolutionary hierarchical dirichlet processes for multiple correlated time-varying corpora,, Association for Computing Machinery, New York, NY, USA.

Zhu, X., Ghahramani, Z., and Lafferty, J. (2005), *Time-sensitive Dirichlet process mixture models* Carnegie Mellon University. Center for Automated Learning and Discovery.