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: February 12, 2025
First draft: July 7, 2020

Abstract

Evolutionary societal changes often prompt a debate. The positions of the two major political parties in the United States on civil rights issues underwent a reversal in the 20th century. The conventional view holds that this shift was a structural break in the 1960s, driven by party elites, while recent studies argue that the change was a more gradual process that began as early as the 1930s, driven by local rank-and-file party members. Motivated by this controversy, this paper develops a nonparametric Bayesian model that incorporates a hidden Markov model into the Dirichlet process mixture model. A distinctive feature of the proposed approach is that it models a process in which multiple latent clusters emerge and diminish as a continuing process so that it uncovers any of steady, sudden, and repeated shifts in analyzing longitudinal data. Our model estimates each party's positions on civil rights in each state based on the legislative activities of their Congressional members, identifying cross- and within-party coalitions over time. We find evidence of gradual racial realignment in the 20th century, with two periods of fast changes during the 1948 election and the Civil Rights Movement.

*We thank Eric Schickler for generously providing the data used in our analysis. We also 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 in 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 remain obscure while they are occurring but become clear after a few decades—often generate 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 Democratic and Republican parties switched their positions on racial issues during the 20th century. At the end of the 20th century, the Democratic Party is associated with racial liberalism, while the Republican Party is linked to racial conservatism. However, the positions of these parties were the opposite in the 1930s. The conventional wisdom often frames this shift as a structural break, in which party elites in Washington, D.C. orchestrated a sudden reshuffling of the party positions in the 1960s (Carmines and Stimson, 1989). Recent studies challenge this view, arguing that the behavior of rank-and-file legislators began to change gradually in the 1940s, or even as early as the 1930s (Schickler, 2016; Schickler et al., 2010; Chen, 2009; Chen et al., 2008).

We revisit this debate by developing a dynamic model of gradual structural changes that extends the Dirichlet process (DP) mixture model (Ferguson, 1973; Antoniak, 1974) to reanalyze historical data. The DP mixture model is a nonparametric Bayesian model for clustering units into latent groups. It places a DP prior on the mixing distribution of a mixture model, allowing the number of clusters to be estimated from data, instead of requiring it to be specified *ex ante*. We extend the DP mixture model to a dynamic setting by modifying the standard DP prior to a Markov process of the DP for changing assignment of units to clusters over time. Like the static DP, this dynamic DP is closely related to the Chinese restaurant process (Blackwell and MacQueen, 1973). We call it the intergenerational Chinese restaurant process because the cluster assignments in each time period are conditioned on the cluster assignments from the previous period.

A distinctive feature of our model is that it captures a process in which multiple clusters emerge and diminish gradually, rather than a one-time structural break. At the same time, our model retains the key feature of the DP mixture model that the number of clusters does not need to be specified *ex ante*. Moreover, as with the DP mixture model, a broad class of models can serve as the mixture component in the dynamic DP mixture model, making it widely applicable beyond our specific focus on party position switching on racial issues.

We use the dynamic DP mixture model to analyze a dataset used by Eric Schickler in his book *Racial Realignment: The Transformation of American Liberalism, 1932-1955*.[1]

[1]This dataset is also featured in Schickler et al. (2010). Schickler also analyzes other types of data, such as survey data and state party platforms, in his book.

This dataset includes four types of legislative activities on civil rights issues—roll-call votes, petition signatures for discharging bills, floor speeches, and bill sponsorships respectively—in the House of Representatives spanning the 73rd to the 92nd Congress (1933—1973).

In our analysis, we treat each party in each state as a unit of analysis, identifying latent clusters among these state-party units based on the legislative activities of their Congressional members. The dynamic DP mixture model enables us to infer gradual changes in their clustering patter over time, thereby answering the question of when and how the two parties in each state changed their positions on racial issues.

Compared to the original analysis by Schickler (2016), our approach offers the advantage of analyzing all types of legislative activities simultaneously, which allows us to leverage a broader set of information and reveal a clearer pattern in how the two parties in each state shifted and realigned their positions. Consistent with Schickler's argument, we observe a distinct trend in the 1940s, in which the two parties in Northern states quickly diverged on civil rights issues. Additionally, we find that since the 1950s, the two parties in each Southern state tended to adopt similar stances on civil rights, a pattern that persisted until the pivotal shifts of the national leaders' positions around 1965 suddenly disrupted this regional alignment between the two parties in the South. We also find that the solidarity of Southern Democrats declined quickly during the Civil Rights Movement.

Our extension of the DP mixture model is characterized by the 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 latent clusters are not shared across time. In this paper, we develop a new dynamic DP prior characterized by "stickiness" of cluster assignments over time (Caron et al., 2012; Fox et al., 2011).

The remainder of the paper proceeds as follows. The next section introduces our motivating empirical example, racial realignment of the Democratic and Republican Parties. 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. We also discuss the differences and similarities between the proposed model and existing models in Section 3. Section 4 presents the empirical analysis of the motivating example, followed by concluding remarks.

2 Motivating Example and Data

2.1 Racial Realignment, 1932-1965

The two major parties in the United States swapped their positions on racial issues with each other during the 20th century. As "Lincoln's party", the Republican Party used to be more closely associated with African Americans than the Democrats at the beginning of the century. By the end of the century, however, the Democratic Party became associated with racial liberalism, advocating for government efforts to address racial inequality, while the Republicans became linked to racial conservatism, more likely to oppose governmental interventions in racial inequality.

Traditionally, scholars view the reversal in the two parties' positions on racial issues as a sudden structural break that occurred in the 1960s. According to this view, although local Northern Democrats had an incentive to support civil rights to gain African American voters, who were becoming increasingly important in the North, federalism constrained local politicians from committing to programmatic liberalism (Weir, 2005). Instead, party elites in Washington, D.C. led the change. The "critical juncture" arrived during the 1964 presidential election, when Democratic candidate Lyndon B. Johnson and Republican candidate Barry Goldwater took sharply opposing positions on civil rights (Carmines and Stimson, 1989; Califano, 1992; Edsall and Edsall, 1992). Subsequently, local party activists aligned their racial positions with those of the national leaders.

By contrast, Schickler (2016) argues that Northern Democrats at the local level took the lead in driving this change.[2] Local Northern Democrats were gradually transformed by the Democratic Party's New Deal coalition with the Congress of Industrial Organizations (CIO), African Americans, and other urban liberals in the North at the local level. Initially, the New Deal coalition, built on shared interests in economic liberalism, had little connection to civil rights advocacy. However, as the Democrats' nonpartisan allies increased their civil rights advocacy, Northern rank-and-file party members followed suit, advancing civil rights advocacy accordingly. While party elites in Washington, D.C. had a strong interest in maintaining the traditional North-South coalition within the Democratic Party, pressures from Northern rank-and-file members eventually forced national leaders to break that solidarity. Meanwhile, the Republican Party, which had previously been more supportive of civil rights, became increasingly divided over racial issues as the demand for civil rights waned in some of its local constituencies.

Schickler (2016) provides extensive data to support his argument, including historical

[2]Following 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.

survey data tracing the configuration of economic and racial liberalism at the individual level, records of state party platforms to identify when and where civil rights issues became (or ceased to be) significant concerns for the two parties at the state level, and data on House members' legislative activities to trace how rank-and-file Congress members—who were more responsible to their state-level constituencies than to party elites in Washington, D.C.—adjusted their racial positions to align with their constituencies' preferences.

In this paper, we reanalyze the legislative data mentioned above to understand how the state-level Democratic and Republican Parties, as represented by their Congress members, broke and remade coalitions on civil rights issues across the North-South divide as well as across party lines, which is a crucial dynamics discussed in Schickler (2016) but never directly analyzed.

2.2 Data

The legislative data includes four types of activities in the House of Representatives: roll-call votes, signing petitions to discharge bills from committee, floor speeches, and bill sponsorships. Roll-call votes are the most commonly used data to measure Congress members' preferences (Poole and Rosenthal, 1985; Clinton et al., 2004). However, Schickler points out that roll-call data can be misleading because only a very small fraction of civil rights bills reached the floor—most were blocked by the House Rules Committee, which was dominated by senior southern Democrats (Schickler, 2016, p. 180-181). To address this limitation, Schickler uses data on petitions, speeches, and bill sponsorships to complement the roll-call measure.

Summary statistics for these four types of legislative data, covering the 73rd Congresses (1933-35) to the 92nd Congress (1971-73), are provided in Supplementary Information (SI) A.

Roll-call Votes. The roll-call vote data consists of a series of dummy variables indicating whether a Congress member supported a civil rights bill: 1 for voting "yes" and 0 for voting "no" or being absent. There may be multiple roll calls on civil rights issues in a given Congress or none in another.

Discharge Petitions. If a legislative committee or the Rules Committee blocks a bill from reaching the floor for a long time, House members can sign a petition to advance the bill.[3] The discharge petition data consists of a series of dummy variables indicating whether a Representative signed a particular petition to discharge a civil rights bill. Similar to roll-

[3]For 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). Schickler explains that signing a discharge petition is costly for members of Congress, because 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 (Schickler, 2016, p. 183).

call votes, there may be no civil rights-related petitions or multiple such petitions in a given Congress.

Floor Speeches. Floor speeches are an important way for Congress members to signal commitments to their constituents. Competition for speech time on the House floor is intense, and using this limited time to deliver a speech in support of civil rights indicates that the issue was among the member's top priorities. The data on floor speeches consists of dummy variables indicating whether a Representative delivered at least one speech on the floor in support of civil rights during a given Congress.

Bill Sponsorship. Like floor speeches, bill sponsorship serves as a signal of strong support for a particular issue. The data on bill sponsorships consists of count variables indicating the number of civil rights bills a Representative sponsored in a given Congress.

2.3 Limitations of the Original Analysis

Schickler (2016) analyzes the four types of data separately.[4] While the analysis of roll-call votes shows that Northern Democrats were only slightly more likely than Northern Republicans to cast votes in support of civil rights before the 1960s, the analysis of discharge petitions reveals that, as early as the 1940s, Northern Democrats were substantially more likely to sign civil rights-related petitions than their Republican counterparts. The newly discovered discharge petition data contribute significantly to the reexamination of racial realignment in the early twentieth century. Previously, scholars relied on roll-call votes to trace changes in Congress members' positions on civil rights. By using discharge petitions for measuring Congress members' preferences on civil rights issues, Schickler (2016) concludes that racial realignment began much earlier than the 1960s. The analysis of speech data and bill sponsorship data further supports this argument.

Analyzing the four types of data separately does not utilize the available information efficiently, and outright disregarding roll-call vote data risks introducing bias in favor of the book's argument. Instead, our model is flexible enough to analyze all four types of data simultaneously, allowing us to make full use of the observed legislative activities to infer the latent preferences of Congress members.

Moreover, Schickler (2016) discusses at length how rank-and-file party members allied across states and party lines in Congress but rarely directly evaluates this argument using data on legislative activities. The original analysis focuses primarily on assessing the claim that Northern Democrats became more supportive of civil rights than the Republican Party well before the 1960s by examining differences in the four types of legislative activities between two parties in the North. However, it does not address other arguments about

[4]Schickler et al. (2010) also presents the results of analyzing the same four types of data.

coalition changes, such as the declining North-South coalition within the Democratic Party or the increasing cross-party coalition between Republicans and Southern Democrats.

Our model addresses the coalition question directly by tracing the evolving clustering patterns of state parties, as represented by their Congress members, over time. This approach allows us to examine how the Democratic Party's North-South coalition dissolved and how the cross-party coalition between Republicans and Southern Democrats emerged.

3 Model and Inference

3.1 The Model

The proposed dynamic DP mixture model is a nonparametric, dynamic Bayesian clustering model. Like the DP mixture model, the dynamic DP mixture model identifies the latent cluster membership of units and is nonparametric in the sense that the number of latent clusters is not specified *ex ante*. However, while the DP mixture model clusters cross-sectional units, the dynamic DP mixture model is a model in which units move across clusters over time. It is a dynamic model since each unit's cluster assignment can change over time, and the temporal shift is modeled as a Markov process. In each time period, a unit either remains in the same cluster as in the previous period or transitions to a different cluster. If a unit leaves its current cluster, the new cluster assignment is determined by a Dirichlet process conditional on the structure of cluster memberships in the previous period. Therefore, the proposed model is a dynamic extension of the DP mixture model.

To formally describe the model, let $i \in \{1, \ldots, N\}$ denote a unit and $t \in \{1, \ldots, T\}$ denote a time period. Also, let Y_{it} be observed (possibly multivariate) measurement for unit i in time t. It is assumed that unit i at time t belongs to a latent cluster $g[it]$, where $g[it] \in \{1, 2, \ldots\}$ denotes the latent cluster index of unit i at time t, and that Y_{it} is generated from a probabilistic model f with parameter $\Theta_{g[it]}$:

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

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

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

In this example, $\Theta_g = (\boldsymbol{\beta}_g, \sigma_g)$ and $\boldsymbol{\beta}_g$ and σ_g vary across latent clusters.

Now we turn to cluster assignments. First, cluster assignments in the initial period

($t = 1$) are generated by a regular DP. Using the stick-breaking construction (Sethuraman, 1994), the generative process of $g[i1], i = 1, \ldots, N$, is hierarchically defined as:

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

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

$$\pi_k \overset{\text{i.i.d.}}{\sim} \mathsf{Beta}(1, \gamma) \tag{4}$$

where γ is the concentration parameter of the DP.

The key innovation of the proposed dynamic DP mixture model is the dynamic generative process of latent clusters for periods $t = 2, \ldots, T$. Specifically, $g[it]$ follows a Markov process conditional on $(g[1, t-1], \ldots, g[N, t-1])$. On the one hand, unit i in time t stays in the same cluster as it was in time $t-1$ with probability p, which we call the stickiness parameter (cf. Fox et al., 2011). Formally,

$$\Pr\left(g[it] = g[i, t-1] \mid g[i, t-1]\right) = p \tag{5}$$

where

$$p \sim \mathsf{Beta}(\alpha_p, \beta_p). \tag{6}$$

On the other hand, with probability $1-p$, unit i's cluster assignment at time t follows a DP, but this DP is conditional on the cluster assignments at $t-1$. That is,

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

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

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

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

3.2 Intergenerational Chinese Restaurant Process Metaphor

The Dirichlet process is known to have two constructive representations: the stick-breaking process and the Chinese restaurant process. Our dynamic DP mixture model, described above, is based on the stick-breaking process due to its mathematical simplicity. However,

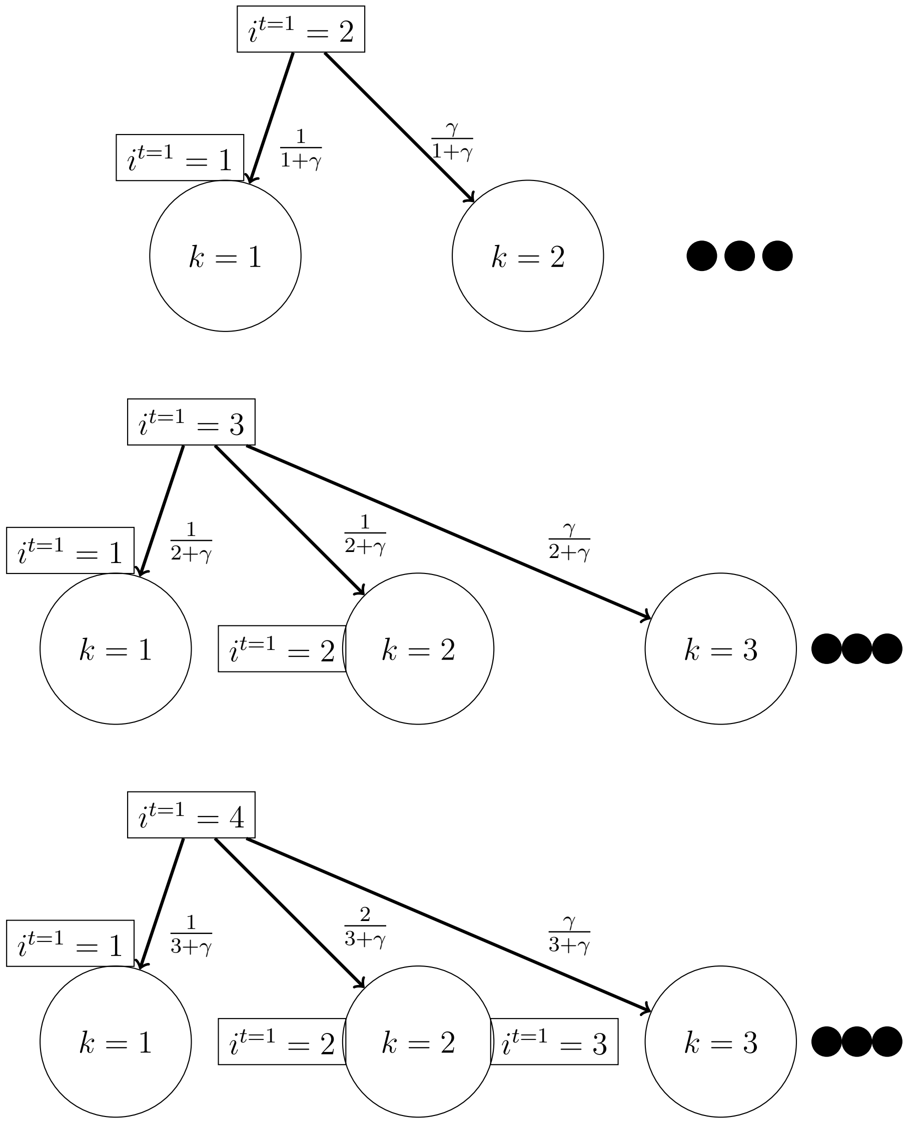


Figure 1: **Chinese Restaurant Process.** This figure illustrates the generative process of latent clusters at $t = 1$. For illustration purposes, we assume that in the realized states, unit 1 is assigned to cluster 1, unit 2 to cluster 2, and unit 3 to cluster 2. The fractions near the arrows represent the probabilities of units being assigned to the corresponding clusters. The probability of assigning a unit to an existing cluster is proportional to the number of units already assigned to that cluster. The probability of a unit creating a new cluster is proportional to the parameter γ. The top, middle, and bottom panels show the probabilities that units 2, 3, and 4 are assigned to each cluster, respectively.

the Chinese restaurant process representation offers a more intuitive illustration of the DP. To further explain our model, we introduce a dynamic variation of the Chinese restaurant process, which we call the intergenerational Chinese restaurant process.

Our dynamic DP mixture model at $t = 1$ can be represented as a regular Chinese restaurant process, which is illustrated in Figure 1. In this representation, units are assigned to latent clusters sequentially. At the beginning, unit 1 is (arbitrarily) assigned to cluster 1. For each subsequent unit, the probability that the unit is assigned to an existing cluster is proportional to the number of the units already assigned to that cluster, and the probability that the unit creates a new cluster is proportional to the concentration parameter γ. For example, unit 2 goes to cluster 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 cluster 2 in the realized state, the next unit (unit 3) is assigned to cluster 1 or 2 with probability $1/(2+\gamma)$, or forms a new cluster (cluster 3) with probability $\gamma/(2+\gamma)$ (the middle panel of Figure 1). Furthermore, given that unit 1 is assigned to cluster 1 and units 2 and 3 are assigned to cluster 2, unit 4 is assigned to cluster 1 with probability $1/(3+\gamma)$, cluster 2 with probability $2/(3+\gamma)$, and cluster 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 clusters at $t = 1$.

The dynamic processes at $t = 2, 3, ..., T$, defined by equations (5) to (9), can be represented as a Chinese restaurant process with added stickiness, which is defined by the stickiness parameter p . Figure 2 illustrates the process at $t = 2$.

The top panel of Figure 2 illustrates the cluster assignment process of unit $1(i^{t=2} = 1)$, the first unit at $t = 2$. For illustration purposes, we assume that in the previous generation of $t = 1$, there are two units in cluster 1 ($i^{t=1} = 1$ and $i^{t=1} = 4$) and two units in cluster 2 ($i^{t=1} = 2$ and $i^{t=1} = 3$). The stickiness between the two generations of $t = 1$ and $t = 2$ is defined by the parameter p. With probability p, unit 1 at $t = 2$ is assigned directly to cluster 1, where the previous generation of unit 1 at $t = 1$ stays; with probability $1 - p$, unit 1 does not go directly to cluster 1 but instead receives a cluster assignment through a Chinese restaurant process, in which the number of units in an existing cluster is determined by both the current generation of $t = 2$ and the previous generation of $t = 1$. If unit 1 does not go directly to cluster 1, it can still be assigned to cluster 1 with probability $2/(4+\gamma)$. Therefore, the total probability that unit 1 at $t = 1$ is assigned to cluster 1 is $p + (1 - p) \times (2/(4 + \gamma))$. The probability that unit 1 is assigned to cluster 2 equals to $(1-p)$ multiplies the conditional probability decided by the Chinese restaurant process, $2/(4 + \gamma)$. Similarly, the probability that unit 1 is assigned to a new cluster is $(1 - p) \times 2/(4 + \gamma)$.

The bottom panel of Figure 2 illustrates the cluster assignment of unit 2 at $t = 2$ ($i^{t=2} = 2$), assuming that unit 1 at $t = 1$ is assigned to cluster 2. Again, unit 2 may directly go to cluster 2, which unit 2 at $t = 1$ stays, or goes through a Chinese restaurant process. Altogether, the probability that unit 2 goes to cluster 2 is $p + (1 - p) \times 3/(5 + \gamma)$. The probability of unit 2 going to cluster 1 is $(1 - p) \times 2/(5 + \gamma)$, and the probability of unit 2 going to a new cluster is $(1 - p) \times \gamma/(5 + \gamma)$.

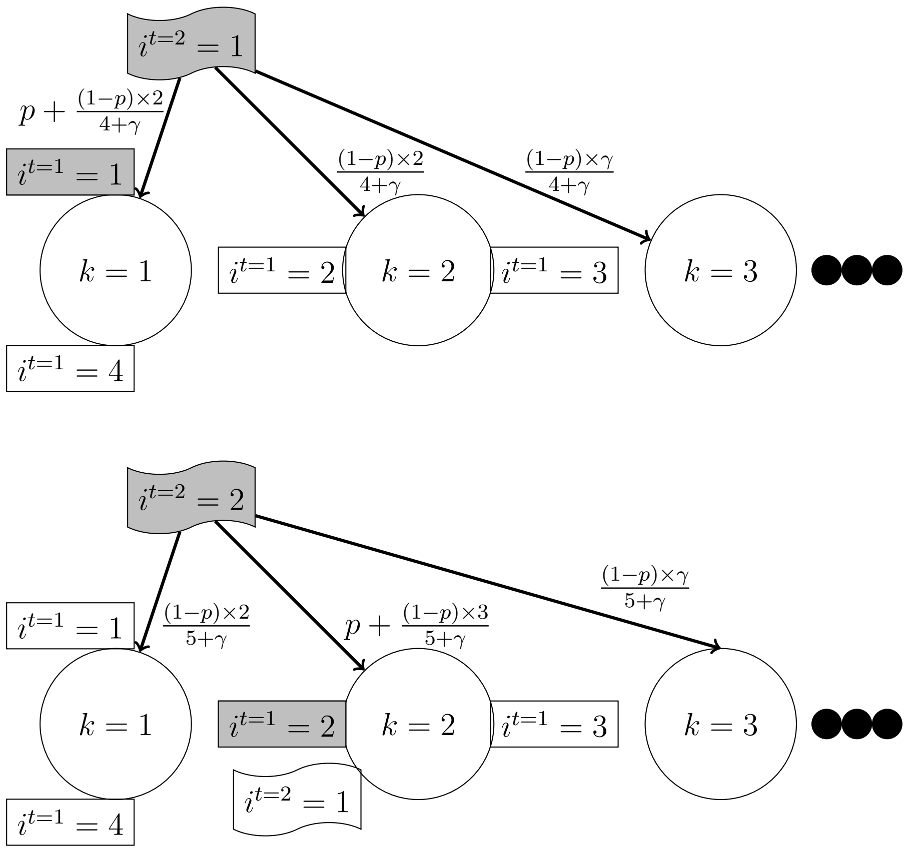


Figure 2: **Intergenerational Chinese Restaurant Process.** This figure illustrates the generative process of the latent clusters at $t = 2$. The fractions near arrows represent the probabilities of units being assigned to the corresponding clusters. The top and bottom panels illustrate the assignment process for unit 1 and unit 2 at $t = 2$, respectively. For illustration purposes, we assume that in the previous generation of $t = 1$, there are two units in cluster 1 ($i^{t=1} = 1$ and $i^{t=1} = 4$) and two units in cluster 2 ($i^{t=1} = 2$ and $i^{t=1} = 3$). In the bottom panel, we assume that unit 1 at $t = 2$ goes to cluster 2. The stickiness between the two generations of $t = 1$ and $t = 2$ is defined by the parameter p.

3.3 Related Models

Among models introducing time dynamics to DP mixture models, the proposed dynamic DP mixture model in our paper is most closely related to Caron et al. (2012). Both models introduce a stickiness parameter. However, in our model, cluster assignments at t depend on the assignments of all units at $t-1$, while in Caron et al. (2012), cluster assignments at t are determined by a selected subset of units at $t-1$, as influenced by the stickiness parameter. This difference results in stronger temporal dependence between generations in our model. For example, in the case of $p = 0$ (i.e., when the cluster assignment of unit i at t is not directly influenced by its assignment at $t = 1$), the cluster assignment process at t is still influenced by cluster assignments of all units at $t-1$. In Caron et al. (2012), no such influence exists in this case.

Iorio et al. (2023) develops another dynamic DP model, in which π's in Equation (8) follow an autoregressive process. In their model, cluster assignments do not depend on the assignments from the previous period. Additionally, the cluster-specific parameters are not shared across time. In Huang et al. (2015), units remain in the same cluster and exit from the dataset 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. In some other models, the DP in period t is a mixture of the DP at $t-1$ and an entirely new DP (Dunson, 2006; Ren et al., 2010; Zhang et al., 2010; Das et al., 2021). In other words, units at t go into the same Chinese restaurant process as $t-1$ or a brand-new Chinese restaurant process.

3.4 Markov Chain Monte Carlo Algorithm for Estimation

In this section, we introduce the Markov Chain Monte Carlo (MCMC) algorithm used to estimate our model constructed through the dynamic stick-breaking process. Specifically, we employ the blocked Gibbs sampling algorithm, which consists of two major steps.

One step is to sample the cluster-specific parameter Θ_g, which defines the data generating function $f(\Theta_k)$ for $k = 1, 2, ...,$ conditioned on the current cluster assignment $g[it] = k$. Sampling Θ_k from its posterior distribution is straightforward when the prior distribution for Θ_k is conjugate.

The second step is to sample cluster assignments $g[it]$ for $i = 1, 2, ..., N$ and $t = 1, 2, ..., T$, conditioned on Θ_k for $k = 1, 2, ...$ Sampling $g[it]$ is more complex, as the posterior distribution is conditioned on both cross-sectional and temporal relationships of unit i at time t with other units. To address this complexity, We combine the forward-backward algorithm designed for change-point models (Chib, 1998) with the truncation approximation approach developed for stick-breaking priors (Ishwaran and James, 2001) to sample $g[it]$.

Following the truncation approximation approach of Ishwaran and James (2001), we begin the algorithm by setting an arbitrarily large number K at which we truncate the theoretically infinite number of clusters. In the posterior distributions of $g[it]$ for all $i = 1, 2, ..., N$ and $t = 1, 2, ..., T$, the probabilities for most clusters will be zero. This approach allows the model to estimate the number of clusters automatically. With K set, we initialize the starting values of $g[it]$ for all $i = 1, ..., N$ and $t = 1, ..., T$.

Each iteration of the Gibbs sampler then proceeds as follows. The first step is to update Θ_k for $k = 1, 2, ..., K$. The posterior distribution of Θ_k is only conditioned on observed data Y_{it} (and X_{it} in the linear regression model example) for units in cluster k. The specific sampling algorithm for Θ_k depends on the form of function $f(\Theta_{g[it]})$.

Second, we sample the stickiness parameter p, which represents the probability that a unit goes directly to the same cluster as in the previous period. To simplify the sampling of p, we introduce augmented binary variables d_{it} for $i = 1, 2, ..., N$ and $t = 1, 2, 3, ...T$. These variables indicate whether a specific unit is directly assigned to the same cluster as before. The posterior distribution of d_{it} is only conditioned on $g[it]$, $g[i, t-1]$, and p. Specifically, When $g[it] \neq g[i, t-1]$, d_{it} must be 0; when $g[it] = g[i, t-1]$, unit i may be assigned directly to the same cluster as before or assigned through a regular DP. Formally,

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

where q_k^t is sampled in the previous MCMC iteration (described below).

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

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

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

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

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

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

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

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

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

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

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

The equation above shows the decomposition of the conditional posterior distribution of $g[it]$ given the observed data, model parameters, and cluster assignments for period $t+1$ through T. Part (1) represents the probability of $g[i, t+1]$ given that $g[i, t]$ is in cluster k. Specifically,

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

Part (2) represents 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:

$$Pr(g[it] = k | p, \mathbf{q}^t, ..., \mathbf{q}^1, \mathbf{Y}_t, ..., \mathbf{Y}_1, \Theta_1, ..., \Theta_K))$$

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

Part (a) is the distribution of observed data Y_{it} conditioned on the cluster assignment $g[it]$, observed data from all the previous periods, and cluster-specific parameters Θ_k for $k = 1, 2, ..., K$. Given $g[it] = k$ and Θ_k, the distribution of Y_{it} is conditionally independent of observed data from the previous periods:

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

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:

$$Pr(g[it] = k | p, \mathbf{q}^t, ..., \mathbf{q}^1, \mathbf{Y}_{t-1}, ..., \mathbf{Y}_1, \Theta_1, ..., \Theta_K)$$

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

There is a 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), while the backward sampling algorithm generates MCMC draws from the conditional posterior of the cluster assignments.

4 Empirical Application

In this section, we apply our proposed dynamic DP mixture model to analyze racial realignment in the twentieth century, using data on legislative activities from Schickler (2016). We begin by explaining how our model is tailored to fit the specific data used in the racial realignment case. Next, we describe the details of the MCMC algorithm employed for estimation. Finally, we present the results of our analysis.

4.1 Statistical Model for Analyzing Racial Realignment

The aim of applying our proposed model to the case of racial realignment is to analyze how state parties allied with each other. We assume that a state party's position on civil rights issues is reflected in its Congress members' legislative activities. Although some Congress members may engage in activities that go against their party's general consensus at the state level, we regard these situations as idiosyncratic. Fundamentally, we assume that Congress members' legislative activities are largely constrained by their state parties. Based on this assumption, we can infer the latent clustering patterns of state parties and how these patterns changed over time using data on legislative activities in the House of Representatives.

Let $t = 73, 74, ..., 92$ index the 73rd Congress (1933-35) to the 92nd Congress (1971-73) and let $s = 1, 2, ..., 50$ index the 50 states. Let D and R represent the Democratic Party and the Republican Party, respectively. Let $i = 1, 2, ..., I_{st}^D$(or I_{st}^R) represent a Democratic (or Republican) Congress member from state s in the t-th Congress, where I_{st}^D is the total number of Democrats in the House of Representatives from State s in the t-th Congress, and I_{st}^R represents the total number of Republicans correspondingly. Let $j = 1, 2, ..., J_t^V$(or J_t^P) index a civil rights vote call or petition in the t-th Congress, where J_t^V represents the total number of civil rights vote calls, and J_t^P the total number of petitions correspondingly.

Let $g^D[st]$ represent the cluster which the Democratic Party of state s in the t-th Congress belongs. Similarly, let $g^R[st]$ represent the cluster of the Republican party. We assume $g^D[st]$ and $g^R[st]$ share the same dynamic DP prior as as defined in Section 3. Here we explain the specific statistical model required in Equation (2) for modeling observed data on legislative activities.

Roll-call Votes. Let V^D_{istj} represent roll-call vote j in the t-th Congress for Democratic member i from state s. Similarly, let V^R_{istj} represent the roll-call vote for a Republican member. We assume V^D_{istj} and V^R_{istj} follow the following distributions:

$$V^D_{istj} \overset{ind.}{\sim} \mathsf{Bernoulli}(\theta_{g^D[st]})$$

$$V^R_{istj} \overset{ind.}{\sim} \mathsf{Bernoulli}(\theta_{g^R[st]})$$

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

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

Discharge Petitions Let P^D_{istj} and P^R_{istj} represent whether member i from state s signed petition j in the t-th Congress, where D and R index Democrats and Republicans, respectively. We assume that P^D_{istj} and P^R_{istj} follow the following distributions:

$$P^D_{istj} \overset{ind.}{\sim} \mathsf{Bernoulli}(\eta_{g^D[st]})$$

$$P^R_{istj} \overset{ind.}{\sim} \mathsf{Bernoulli}(\eta_{g^R[st]})$$

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

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

Floor Speeches. Let S^D_{ist} and S^R_{ist} represent whether Democratic or Republican member i from state s delivered at least one floor speech in support of civil rights during the t-the Congress. We assume that S^D_{ist} and S^R_{ist} follow the following distributions:

$$S^D_{ist} \overset{ind.}{\sim} \mathsf{Bernoulli}(\omega_{g^D[st]})$$

$$S^R_{ist} \overset{ind.}{\sim} \mathsf{Bernoulli}(\omega_{g^R[st]})$$

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

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

Bill Sponsorship. Finally, let B^D_{ist} and B^R_{ist} represent the number of civil rights bills sponsored by Democratic or Republican member i from state s in the t-th Congress. We assume B^D_{ist} and B^R_{ist} follow the following distributions:

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

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

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

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

4.2 MCMC Algorithm

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

4.3 Empirical Results

The MCMC outputs of our model provide the cluster assignments of all state parties in each congressional session. With this information, we can analyze racial alignment patterns across the North-South division, along party lines, or both, and examine how these patterns change over time.

We first investigate whether our model replicates the findings of Schickler (2016), which suggest that in the North, the Democratic Party and the Republican Party began to diverge in their civil rights positions during the 1940s. To achieve this, we analyze whether and when, in the North, the Democratic Party and the Republican Party at the state level were less likely to belong to the same cluster. We perform a similar analysis for the South to examine the pattern of cross-party alignment and how it changes over time.

Next, we test two arguments frequently mentioned by Schickler (2016) but not directly tested with legislative data. The first argument is that the Democratic Party's North-South coalition dissolved long before the 1960s. The second argument is that the Republican Party in the North became increasingly divided over civil rights issues before the 1960s. We test these arguments by analyzing the following: (1) Within the Democratic Party, whether and when northern state parties were less likely to be in the same cluster as southern state parties; (2) Within the Northern Republican Party, whether and when northern state parties were less likely to be in the same cluster.

Finally, we examine the solidarity of the Democratic Party in the South—an assumption implicitly made by Schickler (2016) but not tested—by analyzing how the probability of two Southern states being in the same cluster within the Democratic Party changes over time.

Cross-party Alignment. The patterns of cross-party alignment in the North and South respectively are presented in Figure 3. Panel (a) shows the pooled results: each point represents the average probability that a state Democratic Party and a state Republican Party are in the same cluster in the North (or South) for the corresponding Congress, with a 95% credible interval. Panel (b) shows the within-state results: each point represents the average probability that, within a given state, the Democratic Party and the Republican Party are in the same cluster for the corresponding Congress, with a 95% credible interval.

The pooled results and within-state results show consistent patterns of cross-party alignments. First, the probabilities that the two parties in the North are in the same cluster decline rapidly in the 1940s. This replicates the findings in Schickler (2016), which indicate that by the end of the 1940s, the two parties in the North had diverged in their positions on civil rights.

Panel (a) shows that toward the end of the 1930s, northern Democrats are more likely to be in the same cluster as northern Republicans. This aligns with Schickler's finding that after the New Deal, northern Democrats quickly caught up with Republicans in supporting civil rights.

Panel (b) shows that Democratic and Republican parties within the same Northern state have already had a high probability of acting together even before the New Deal.

In addition to analyzing cross-party alignment in the North, as discussed in Schickler (2016), we also perform the same analysis for Southern states. Interestingly, both pooled and within-state results show that in early 1950s, the probabilities that the two parties in the South are in the same cluster increased rapidly. Previously, the probabilities were very low (close to 0).

This result contradicts Schickler's claim that Southern Republicans began aligning with racially conservative Southern Democrats only after conservatives within the GOP captured the southern party machinery established by President Dwight Eisenhower (Chapter 10). Our results show that, even before Eisenhower's party-building efforts, Southern Republicans had already aligned with Southern Democrats.

Finally, while panel (a) explores the general pattern of cross-party alignment over civil rights issues in the North and South, panel (b) examines the within-state alignment pattern. This allows us to investigate whether, when facing the same socioeconomic conditions, Democrats and Republicans are more likely to align with each other on civil rights issues. In general, the probabilities of the two parties in the same cluster are higher in panel (b) than in panel (a), indicating that within the same state, the two parties are more likely to take similar positions on civil rights issues.

Alignment with Southern Democrats. We now turn to evaluate an argument frequently made by Schickler (2016), but not tested directly, that the Democratic Party's North-South coalition declined long before the 1960s. In Figure 4, we calculate the average probability that a northern Democratic Party is in the same cluster as a southern Democratic Party in each year. For comparison, we also calculate the average probability that a northern Republican Party is in the same cluster as a southern Democratic Party, since Schickler argues that the GOP in some northern states, particularly in the rural Midwest, had already been more likely to align with southern Democrats before the 1960s.

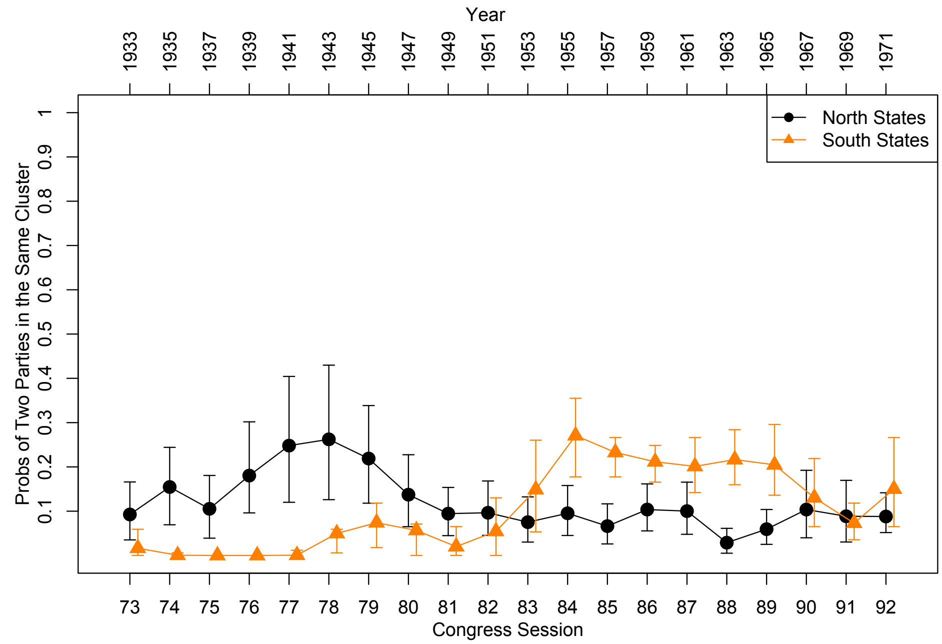
Figure 4 shows that, consistent with Schickler's argument, the probability that northern and southern Democrats are in the same cluster declines rapidly in the 1930s. By the early 1940s, northern Democrats, like northern Republicans, have almost no chance of being in the same cluster as southern Democrats.

Toward the late 1940s and early 1950s, both northern Democrats and northern Republicans suddenly became more likely to align with southern Democrats. However, compared to northern Democrats, northern Republicans had a higher probability of being in the same cluster as southern Democrats during this period of peak alignment. After this period of sudden change, the probabilities of northern Republicans and southern Democrats being in the same cluster remain higher than those of northern and southern Democrats being in the same cluster.

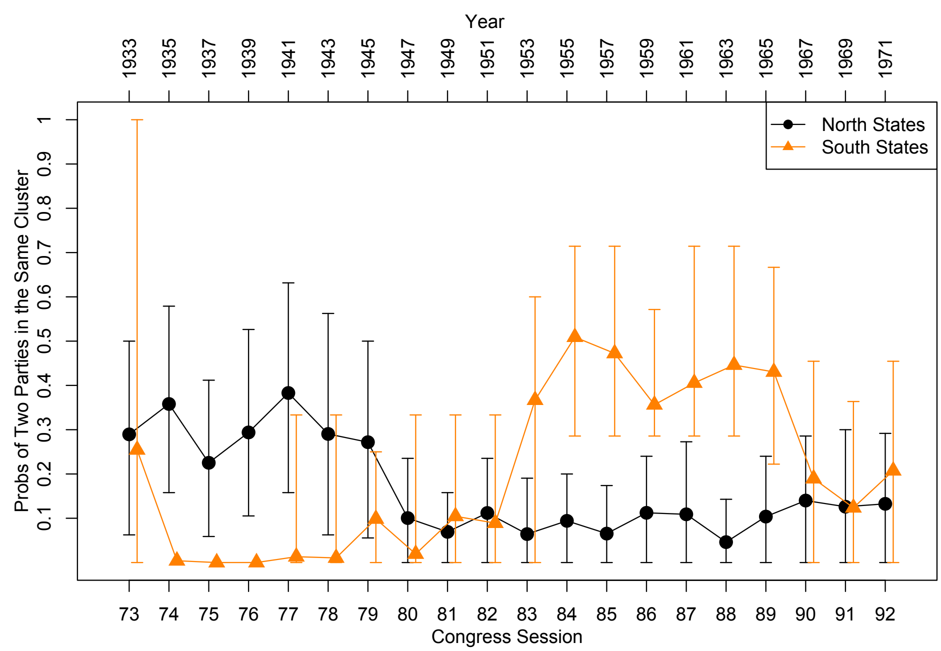
Within-Party Solidarity: Northern Republicans and Southern Democrats. Finally, we examine another untested argument in Schickler (2016) that Northern Republicans had already become divided in the 1940s and 1950s over civil rights issues. Additionally, we evaluate the solidarity of Southern Democrats—an assumption made by Schickler (2016) but not formally tested. In Figure 5, we calculate the average probability that a northern Republican Party in one state is in the same cluster as the Republican Party in another northern state in each year. Similarly, we calculate the average probability that a southern Democratic Party in one state is in the same cluster as the Democratic Party in another southern state in each year.

Consistent with Schickler (2016), Figure 5 shows that the probability of two northern Republican parties being in the same cluster declines steadily in the 1940s and 1950s. However, in contrast to Schickler's assumption that Southern Democrats were a solid unity against

Figure 3: Cross-party Alignment, in the North and South



(a) **Pooled Results.** Each point represents the average probability that a state Democratic Party and a state Republican Party are in the same cluster in the North (or South) for the corresponding Congress, with a 95% credible interval.



(b) **Within-state Results.** Each point represents the average probability that, within a given state, the Democratic Party and the Republican Party are in the same cluster for the corresponding Congress, with a 95% credible interval.

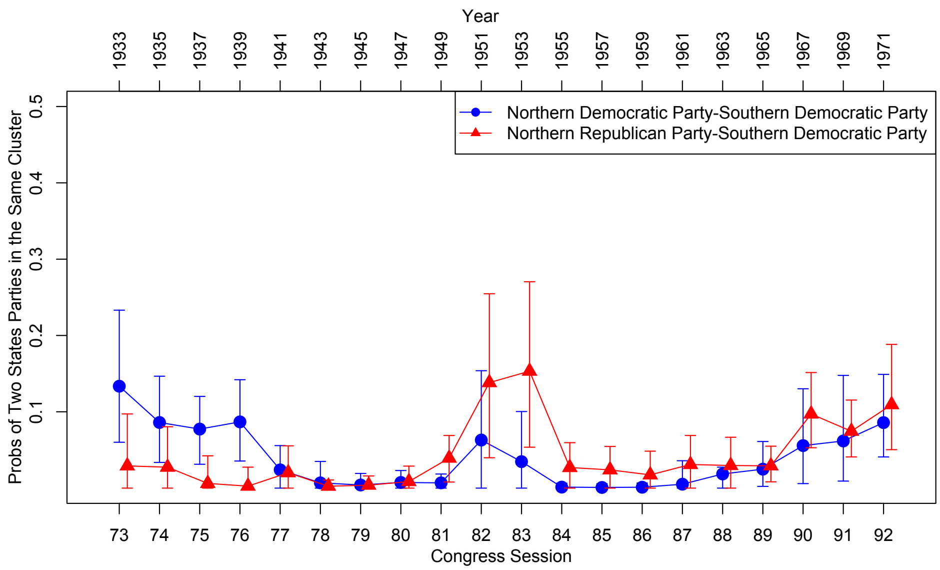


Figure 4: **Alignment of Southern Democrats with Northern Democrats and Republicans Respectively.** Each point represents the average probability that a state Democratic Party (or Republican Party) in the North and a state Democratic Party in the South are in the same cluster, with a 95% credible interval.

civil rights, their regional coalition peaked in the early 1950s and began to decline rapidly after 1955, during the tumultuous years of the civil rights movement.

5 Concluding Remarks

Most societal changes are evolutionary rather than revolutionary, and it is difficult to capture gradual changes as opposed to abrupt ones. Debate over the timing of the racial realignment between two major parties in the United States illustrates this difficulty. Although there was a clearly visible change at the level of party leadership in the 1960s, on which the conventional wisdom is based, statistical analysis of newly available data suggests that the realignment was a gradual process that began in the 1930s.

This paper develops a novel dynamic Dirichlet mixture model for analyzing evolutionary changes. In the proposed model, units are clustered into latent clusters, and cluster assignments are assumed to follow a Markov process. In each time period, a unit remains in the same cluster as the previous period with a probability controlled by a stickiness parameter or moves to another cluster with a probability determined by the Chinese restaurant process, conditional on the cluster assignments in the previous period. Because of this cluster assignment process, the proposed dynamic DP model can effectively capture gradual structural changes.

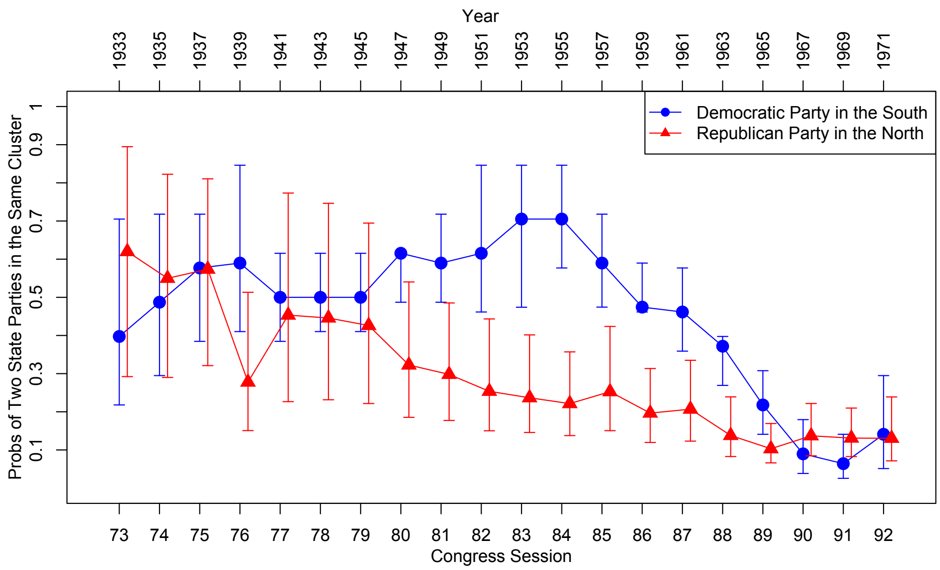


Figure 5: **Within-party Solidarity, Northern Republicans and Southern Democrats.** Each point represents the average probability that two Republican parties in the North (or Democratic parties in the South) are in the same cluster, with a 95% credible interval.

Our empirical analysis of racial realignments demonstrates that the switch in positions on civil rights between the Democratic Party and the Republican Party had been a gradual process since the New Deal, rather than a sudden structural break in the 1960s. We found that the Democratic Party and the Republican Party began to take divergent positions on civil rights issues in the 1940s. Moreover, our analysis identifies two key moments of racial realignment that have not been tested with data in existing studies. First, at the end of the 1940s and the beginning of the 1950s, both Northern and Southern Republicans were more likely to be aligned with Southern Democrats. President Truman's postwar civil rights advocacy during the 1948 election, which triggered the Dixiecrat bolt and deep Southern anger toward the Truman administration, may have facilitated this coalition between Republicans and Southern Democrats. Second, during the Civil Rights Movement, starting in 1954, solidarity within the Southern Democratic Party gradually dissolved.

Beyond the case of racial realignment, the proposed model is widely applicable to any dataset with repeated measurements of multiple units. Its use would therefore benefit many social scientists interested in the temporal dynamics of unobserved heterogeneity.

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.

Carmines, E. G., and Stimson, J. A. (1989), *Issue Evolution: Race and the Transformation of American Politics*, Princeton, NJ: Princeton University Press.

Caron, F., Davy, M., and Doucet, A. (2012), "Generalized Polya Urn for Time-varying Dirichlet Process Mixtures,".

Chen, A. (2009), *The Fifth Freedom: Jobs, Politics, and Civil Rights in the United States, 1941-1972*, Princeton, NJ: Princeton University Press.

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.

Clinton, J., Jackman, S., and Rivers, D. (2004), "The Statistical Analysis of Roll Call Data," *American Political Science Review*, 98(2), 355–370.

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.

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.

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.

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.

Poole, K. T., and Rosenthal, H. (1985), "A Spatial Model for Legislative Roll Call Analysis," *American Journal of Political Science*, pp. 357–384.

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.

Weir, M. (2005), "States, Race, and the Decline of New Deal Liberalism," *Studies in 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,, in *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '10, Association for Computing Machinery, New York, NY, USA, p. 1079–1088.

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