

Discussion Paper Series A No.655

**Estimation of Unobserved Dynamics of
Individual Partisanship: A Bayesian Approach**

Tomohito Okabe
(Institute of Economic Research, Hitotsubashi University)
and
Daisuke Nogiwa
(Department of Management and Information Science,
Fukui University of Technology)

March 2017

Institute of Economic Research
Hitotsubashi University
Kunitachi, Tokyo, 186-8603 Japan

Estimation of Unobserved Dynamics of Individual Partisanship: A Bayesian Approach

Tomohito Okabe* Daisuke Nogiwa†

March 13, 2017

Abstract

Political party preference is a crucial element in the analysis of economics and political science. However, it is often difficult to investigate the dynamic properties of the individual partisanship due to inaccessibility to panel data. This study proposes a Bayesian approach for estimating Markov dynamics of individual-level partisanship with repeated cross-section data in which the history of respondents' choice of favored party cannot be observed. The proposed approach identifies individual heterogeneities that affect transitional patterns of partisanship, and replicates the dynamic patterns of individual partisan mobility. Using the proposed method with American survey data, the study shows that age, education and race significantly influence partisan dynamics among Americans for three decades from 1972.

Keywords: Party Identification; Repeated Cross-Section Data; Markov Chain; Bayesian Microeconometrics

JEL Classification: C25; D39; D72

*Institute of Economic Research, Hitotsubashi University. E-mail address: t-okabe@ier.hit-u.ac.jp.

†Department of Management and Information Science, Fukui University of Technology. E-mail address: d-nogiwa@fukui-ut.ac.jp.

We thank Timothy Kam, John Stachurski and Gaurab Aryal for their advice and encouragement. We are grateful to Takuya Satomura and Eisaku Sato for their helpful suggestions. We appreciate comments from the participants of the RSE Brown-bag Seminar at Australian National University and Japan Institute of Marketing Science Workshop at Keio University.

1 Introduction

Individual partisanship has played an important role in the study of voting behavior in elections. Political scientists have investigated empirically the dynamics of voter preference with respect to political parties (e.g., [Jackson and Kollman, 2011](#)). Such partisan preference is recognized as an important concept in political studies in the US and beyond and is referred to as party identification (PID). Specifically, PID is defined as “an attachment to a party that helps the citizen locate him/herself and others on the political landscape” ([Campbell et al., 1986](#), pg.100). It can be also viewed as a viable proxy for an individual’s position in the ideological space, that is, the conservative-liberal or right-left-wing space.

The most relevant economic theory regarding individual ideology in elections is the probabilistic voting game, in which voters are categorized into specific groups of interests. As different voter groups have conflicting preferences regarding policy variables, conditional on ideology, candidates with distinct ideologies seek to maximize their own vote share by announcing their policy platforms. Equilibrium policies, in turn, hinge on the joint distribution of voter types (i.e., groups of interests) and ideologies. Thus, voter ideologies influence equilibrium policies. This model suggests the manner through which ideology (and its proxy, PID) affects electoral consequences. For more details of the voting model, see e.g., [Persson and Tabellini \(2002\)](#); [Banks and Duggan \(2005\)](#).

Given the relevant partisanship literature in political science and its applicability to economics, the present study seeks to shed light on the dynamic properties of PID. In particular, the long-term transitions in American individual preference with respect to political parties are investigated. A number of empirical studies (e.g. [Clarke and McCutcheon, 2009](#); [Bartels et al., 2011](#)) have examined individual-level partisan movement, but their focus has largely been on the short-term transitions owing to the lack of long-spanning panel data. In individual-level dynamics, most econometric methods require panel data sets covering time periods of interest. However, political science surveys tend to offer limited panel waves, thereby directly restricting researchers’ choice of time periods for analysis. [Bartels et al. \(2011\)](#) noted that the lack of long-period panel data and appropriate estimation methodology impedes future work on PID transitional patterns. Thus, the following question is posed: Is there any way to estimate the individual-level partisanship dynamics from data in which transitions in individual favored parties cannot be traced?

The current work proposes a Bayesian approach for estimating individual-level PID dynamics with repeated cross-section (RCS) data, which are much more easily available than panel data. Specifically, we present an econometric model in which we presume that individual PID changes over time according to a three-state Markov chain. The three states stand for the PID of American peoples: Democrat, Independent and Republican. We label the presented model as Markov chain model. Our Markov chain model characterizes the unobserved PID dynamics of each respondent with his or her time-invariant heterogeneities (e.g. gender, race), thereby allowing for the computation of the unconditional probability distribution for the respondents’ PID in the surveyed period. At this stage, respondents with the same heterogeneities are treated as identical people even if they belong to samples in different periods. We then set up priors for unknown parameters embedded in the Markov chain model, and obtain the posterior distributions using a Markov Chain Monte Carlo (MCMC) method. The MCMC method is employed essentially because the likelihood function for parameterizing unconditional probability distribution is given in a complex non-linear form, which hinders the use of frequentist estimation techniques, including the maximum likelihood method.

The present work estimates transitional patterns of American partisanship, referring to gradual transition theory ([Mebane and Wand, 1997](#)). The theory states that Americans tend to change their partisanship

gradually through the Independent position, rather than jumping across extremes (i.e., from Democrat to Republican or vice versa). In our estimation, this serves as prior knowledge for enriching the limited information on individual-level PID transitions in RCS data. Given the gradual transition theory, our simulation exercise with artificial data shows that our proposed method is rather valid to replicate true Markov dynamics patterns.

Relevant literature includes [Moffitt \(1993\)](#) who produced a seminal work for the estimation of Markov properties with RCS data. [Pelzer et al. \(2002\)](#) extended Moffitt’s analysis using a more flexible framework. Our study shares the basic structure of their Markov models, but differs in two aspects. First, their models are utilized for a two-state dynamics whereas our model is for a three-state dynamics, applicable for a Markov process with more discrete states. Second, their estimation uses the maximum likelihood method because their two-state Markov chain model allows for parameterizing the likelihood function in analytical form. In contrast, we cannot apply the maximum likelihood method and thus estimate with the MCMC method.

We estimate the American partisanship dynamics using American national survey data for three decades from 1972 to 2002. The results of the estimation provide interesting empirical findings. Specifically, race significantly affects Americans’ partisanship. In particular, African Americans tend to remain as the Democrats whereas White Americans do as the Republicans. This finding is consistent with previous claims that African Americans largely become ardent Democrats over time (e.g., [Fiorina and Abrams, 2008](#)). Race affects the dynamics of Independent partisanship as well. We show that age and education also play important roles in characterizing individual partisan dynamics.

The remainder of this paper is organized as follows. Section 2 presents the econometric model, the proposed estimation method and the simulation exercise. Section 3 shows the results of estimation. Finally, Section 4 concludes the paper.

2 Estimation of Transitional Probabilities

2.1 PID Markov Chain

In a three-state space representing American PID $\Omega := \{1, 2, 3\}$, where 1 is a label for *Democrat*, 2 for *Independent* and 3 for *Republican*, the marginal probabilities for individual political position $\mathbf{p}(t) := \{p_1(t), p_2(t), p_3(t)\}$ satisfy the following:

$$p_k(t) > 0 \quad \forall k \in \Omega,$$

and

$$\sum_{k=1}^3 p_k(t) = 1.^1 \tag{1}$$

Suppose that individual PID changes over time, such that the current-period position solely depends on the last-period position. This Markov dynamics is then written as:

$$\mathbf{p}(t) = \mathbf{p}(t - 1)\mathbf{\Lambda}, \tag{2}$$

¹Our analysis is applicable for a Markov chain with more discrete states. The generalization to a higher-order state space is straightforward, and hence omitted here to save space.

where

$$\mathbf{\Lambda} := \begin{bmatrix} \lambda_{1|1} & \lambda_{2|1} & \lambda_{3|1} \\ \lambda_{1|2} & \lambda_{2|2} & \lambda_{3|2} \\ \lambda_{1|3} & \lambda_{2|3} & \lambda_{3|3} \end{bmatrix},$$

with

$$\sum_{\ell=1}^3 \lambda_{\ell|k} = 1 \quad \forall k \in \Omega, \quad (3)$$

and

$$\lambda_{\ell|k} \geq 0 \quad \forall (\ell, k) \text{ in } \Omega \times \Omega. \quad (4)$$

The transitional matrix element $\lambda_{\ell|k}$ is equivalent to the conditional probability that alternative ℓ is realized in the current period, given the last period's alternative k . In our paper, the terms “transitional matrix” and “Markov matrix” are used interchangeably.

2.2 Econometric Model

We deal with RCS data that document individual survey responses on PID and exogenous characteristics (e.g., age, gender, race) for $(\tau + 1)$ time periods. The data do not allow tracing of individual responses over time, unlike in the case of panel data. Table 1 illustrates the RCS data. Our problem is in estimating individual-level dynamics with RCS data. First, we consider parameterizing the Markov matrix with individual characteristics. For notational convenience, we introduce a unit index i in the transitional matrix that should vary across individuals:

$$\mathbf{\Lambda}^{(i)} := \begin{bmatrix} \lambda_{1|1}^{(i)} & \lambda_{2|1}^{(i)} & \lambda_{3|1}^{(i)} \\ \lambda_{1|2}^{(i)} & \lambda_{2|2}^{(i)} & \lambda_{3|2}^{(i)} \\ \lambda_{1|3}^{(i)} & \lambda_{2|3}^{(i)} & \lambda_{3|3}^{(i)} \end{bmatrix} \quad (5)$$

Second, using a multinomial logistic function, we parameterize the conditional probabilities as:

$$\lambda_{\ell|k}^{(i)} = \frac{\exp(\mathbf{x}_i' \boldsymbol{\beta}_{\ell|k})}{\sum_{m=1}^3 \exp(\mathbf{x}_i' \boldsymbol{\beta}_{m|k})}, \quad \text{for } \ell \neq k \quad (6)$$

$$\lambda_{k|k}^{(i)} = \frac{1}{\sum_{m=1}^3 \exp(\mathbf{x}_i' \boldsymbol{\beta}_{m|k})} \quad (7)$$

where \mathbf{x}_i is an M -dimensional covariate vector of time-invariant characteristics of individuals, and $\boldsymbol{\beta}_{\ell|k}$ is a vector of parameters. The benchmark partisan position is the position in the previous period; hence, its coefficients are normalized to zero, i.e., $\boldsymbol{\beta}_{k|k} = \mathbf{0} \quad \forall k \in \Omega$. Notice that this set up imposes no restrictions on the values of $\boldsymbol{\beta}_{\ell|k}$, whereas it satisfies the usual restrictions (1) and (4) for the Markov matrix. In establishing

the multinomial logistic function, we assume that an individual chooses an alternative, given his/ her last period's choice such that it maximizes his/ her utility, conditional on his/ her characteristics. The parameter $\beta_{\ell|k}$, hence, measures the impact of individual characteristics on utility (see appendix A for more details).

The conditional probabilities immediately allow for computing the marginal probabilities at time τ when individual i is observed, given the initial probabilities. Specifically, the marginal probabilities are given by:

$$\mathbf{p}_i := (p_{i,1}, p_{i,2}, p_{i,3}) := \mathbf{p}(0) \left(\mathbf{\Lambda}^{(i)} \right)^\tau \quad (8)$$

Notice that the marginal probabilities are given by a mapping of $\mathbf{\Lambda}^{(i)}$ and hence of \mathbf{x}_i and the whole set of parameters β , i.e., $\mathbf{p}_i = \tilde{\mathbf{p}}_i(\mathbf{x}_i, \beta)$. The likelihood function is then defined as:

$$L(y|\mathbf{x}, \beta) = \prod_{i=1}^I p_{i,1}^{\mathbb{1}_{\{y_i=1\}}} p_{i,2}^{\mathbb{1}_{\{y_i=2\}}} p_{i,3}^{\mathbb{1}_{\{y_i=3\}}} \quad (9)$$

where y is the set of the PID responses y_i , and \mathbf{x} is the whole set of \mathbf{x}_i . $\mathbb{1}_{\{y_i=z\}}$ denotes the indicator function taking the value 1 if $y_i = z$ and 0 otherwise. As the derived likelihood function is complex and non-linear with respect to the parameters β , we shall numerically evaluate the likelihood and hence employ the Bayesian estimation technique rather than the frequentist methods.²

Taking into consideration a standard set up in Bayesian statistical literature (see Rossi et al., 2005), we choose a multivariate normal distribution for the parameter prior with an inverse-Wishart distribution for the covariance prior:

$$\begin{aligned} \beta_{|k} &\sim MVN(\mathbf{0}, \mathbf{V}_\beta \otimes \mathbb{I}_M) \\ \mathbf{V}_\beta &\sim IW(\nu_0, \mathbf{V}_0) \end{aligned}$$

where $\beta_{|k} = (\beta_{1|k}, \beta_{2|k}, \beta_{3|k})$ where $\nu_0 > 0$ is a constant, and \mathbf{V}_0 is a scale matrix. Finally, the Bayes' theorem yields the posterior distributions:³

$$\pi(\beta, \mathbf{V}_\beta | y, \mathbf{x}) \propto L(y|\beta, \mathbf{x}) \pi(\beta | \mathbf{V}_\beta) \pi(\mathbf{V}_\beta) \quad (10)$$

2.3 Embedding Gradual Transition Theory

As our RCS data convey limited information, we utilize a political theory as prior knowledge. Mebane and Wand (1997) asserted that individual Americans PID gradually changes in the short term, always transitioning thorough Independent status rather than jumping from one extreme to the other (i.e., Democrat to Republican and vice versa). We label this the ‘‘Mabane-Wand gradual transition theory’’, and stress that this theory is supported by empirical evidences. For instance, table 2 is the transitional rates of American PID which Bartels et al. (2011) computed from a panel data set.⁴ Observe that the bold numbers for diagonal elements $\lambda_{3|1}, \lambda_{1|3}$ are nearly equal to zero.

²In contrast, a two-state Markov chain model allows for applying the maximum likelihood method because the initial distribution $\mathbf{p}(0)$ is eliminated in the derivation of the marginal probabilities. The marginal probabilities then turn out to be a function of reduced form parameters. See Moffitt (1993) for more details.

³The general Bayes' theorem is given by: $f(\gamma|z) = \frac{f(z|\gamma)f(\gamma)}{f(z)}$, where γ is a vector of parameters and z is a vector of observations. The numerator is the product of a likelihood function and a prior distribution. The denominator is a constant, i.e., $f(z) = \int f(z|\gamma)f(\gamma)d\gamma$. Therefore, $f(\gamma|z) \propto f(z|\gamma)f(\gamma)$.

⁴Their data source is *American National Election Studies* (ANES) panel in 1992 - 1996.

We can readily mimic this dynamic pattern by imposing on our Markov chain model linear restrictions, $\lambda_{3|1}^{(i)} = \lambda_{1|3}^{(i)} = 0$. The conditional probabilities for the first and third row of $\Lambda^{(i)}$ are then recast as:

$$\lambda_{\ell|k}^{(i)} = \frac{\exp(\mathbf{x}'_i \boldsymbol{\beta}_{\ell|k})}{1 + \exp(\mathbf{x}'_i \boldsymbol{\beta}_{\ell|k})}, \quad \text{for } \ell = 2, k \in \{1, 3\} \quad (11)$$

$$\lambda_{k|k}^{(i)} = \frac{1}{1 + \exp(\mathbf{x}'_i \boldsymbol{\beta}_{\ell|k})}, \quad \text{for } \ell = k \in \{1, 3\} \quad (12)$$

Observe that the link function is replaced by a binary logistic function. The priors are then given by:

$$\boldsymbol{\beta}_{\ell|k} \sim N(0, v_0)$$

The gradual transition theory accommodates estimation difficulty by narrowing down possible Markov processes. Recall that we essentially attempt to conjecture an unobservable subsequent transitional path, i.e., from which PID position a respondent comes, given his/ her current PID position. Figure 1 and figure 2 show how the potential paths are confined. In both figures, given an observed PID ℓ , a conditional probability $\lambda_{\ell|}$ is expressed by an inflow arrow directing to the observed PID. The former depicts a plain-vanilla Markov chain, in which each PID has three outflows: $\lambda_{\ell|1}$, $\lambda_{\ell|2}$ and $\lambda_{\ell|3}$. The latter depicts the Markov chain embedded with the gradual transition theory, in which the outflows from Democrat and Republican are reduced to twain: $\lambda_{\ell|\ell}$ and $\lambda_{\ell|2}$. Thus, the prior theory's information narrowing potential paths alleviates uncertainties arising from limited data information, thus facilitating our estimation.

2.4 MCMC Method

We estimate the parameters with a Bayesian approach. Specifically, we perform Hamiltonian Monte Carlo (HMC) sampling. The HMC sampling is an MCMC method that allows for much more efficient sampling than simple methods, such as random-walk Metropolis. Letting $\boldsymbol{\theta} := \{\boldsymbol{\beta}, \mathbf{V}_\beta\}$ denote the set of all parameters, our goal is to draw from a Bayesian posterior $\pi(\boldsymbol{\theta}|y, \mathbf{x})$. Thus, we introduce auxiliary momentum variables $\boldsymbol{\rho}$ that are independent of the parameters, and then draw from a joint density:

$$\pi(\boldsymbol{\rho}, \boldsymbol{\theta}|y, \mathbf{x}) = \pi(\boldsymbol{\rho}) \pi(\boldsymbol{\theta}|y, \mathbf{x})$$

Then, given an initial set of parameters $\boldsymbol{\theta}$, the auxiliary variables $\boldsymbol{\rho}$ and parameters $\boldsymbol{\theta}$ are updated for a given number of iterations. The updating process consists of three stages. First, new momentum $\boldsymbol{\rho}$ and the current value of parameters $\boldsymbol{\theta}$ are evolved according to Hamiltonian dynamics:

$$\begin{cases} \frac{d\boldsymbol{\theta}}{dt} = \frac{\partial H}{\partial \boldsymbol{\rho}} \\ \frac{d\boldsymbol{\rho}}{dt} = -\frac{\partial H}{\partial \boldsymbol{\theta}} \end{cases} \quad (13)$$

where $H := -\log \pi(\boldsymbol{\rho}, \boldsymbol{\theta})$ is the Hamiltonian. Second, the system of differential equation (13) is solved using numerical integration with a small time interval. Finally, the Metropolis accept step is applied. The

acceptance probability for the proposal $(\boldsymbol{\rho}^*, \boldsymbol{\theta}^*)$ generated by transitioning from $(\boldsymbol{\rho}, \boldsymbol{\theta})$ is given by:

$$\min \left(1, \frac{\pi(\boldsymbol{\rho}^*, \boldsymbol{\theta}^*)}{\pi(\boldsymbol{\rho}, \boldsymbol{\theta})} \right)$$

In practice for the HMC sampling, we use open-access statistical software, Stan (<http://mc-stan.org/>). See Neal (2011) and Stan documents (<http://mc-stan.org/>) for more details of the computation algorithm.

2.5 Simulation Study

In this section, we illustrate the validity of our proposed method with artificial panel data which we created using Markov chains. The created data samples include 1,000 individuals with two binary characteristics for 10 observation periods. Individuals are disaggregated into four types by the combinations of dummy variables for characteristics, i.e., 250 identical individuals per type. Individual PID observations are then randomly generated by Markov chains of which transitional matrices differ across types. Table 3 summarizes the definitions of individual types.

In the estimation, we treat the generated panel data as if they are RCS data. The MCMC sampling is conducted with four chains, each of which contains 10,000 iterations with the burn-in period of 5,000. The value of potential scale reduction statistic factor \hat{R} is monitored, and maintained at the value of 1 for all posteriors, which ensures the convergence of computation. See Gelman and Rubin (1992) for the convergence diagnosis.

Table 4 shows the posterior distributions of parameters. We interpret a coefficient estimate as statistically significant if its 95 % credible interval does not include zero. Observe that the other coefficient parameters for the Characteristic B dummy in $\beta_{1|2}$ and $\beta_{3|2}$ are insignificant whereas the others are all significant. The insignificance for the former suggests that Characteristic B does not affect the PID mobility from party 2 in the last period, which is consistent with the true dynamics. Notice that, in table 3, $\lambda_{\cdot|2}$ are identical between Type I and Type II, and between Type III and Type IV.

Table 5 shows the estimated transitional matrices which are computed from the logistic functions (6), (7) and (11) and (12) with the mean values of significant parameters. In the computing process, we set the values of insignificant parameters as zero. It should be stressed that the overall patterns of the reproduced matrices are fairly close to the true values.⁵ This is because the signs of the significant coefficients are consistent with the patterns of true transitional matrices. For instance, in table 4, the mean value for Characteristic A dummy in $\beta_{2|1}$ takes 0.72, which implies that an individual with Characteristics A is more likely to move to party 2, given her last period’s position of party 1 if all else is held constant. In other words, $\lambda_{1|1}$ takes smaller values when the Characteristic A dummy is equal to 1. In fact, in table 5, the estimate (true value) for $\lambda_{1|1}$ for Type I equals 0.66 (0.75) which is smaller than 0.8 (0.8) for Type III. Also, the estimate (true value) for $\lambda_{1|1}$ for Type II equals 0.23 (0.35) which is smaller than 0.38 (0.4) for Type IV. As for other coefficients and PID mobility, we can also establish arguments in a similar manner, but omit them to save space. In sum, the simulation results show that our proposed method is rather valid for capturing distinct differences in partisan dynamics, given a lack of panel data.

⁵An exception is the estimates for $\lambda_{\cdot|3}$ in Type I and Type IV. Although the signs of coefficients are correctly estimated, the mean values of the posteriors are not precise enough to replicate the true values.

3 Results

3.1 Data and Strategy

The employed data is the *American National Election Studies* (ANES) cross-section survey which is a national-level public opinion survey widely used in political science and political economy studies. The cross-section data cover the period from 1972 to 2002 in 16 waves.⁶ PID observations are collected in interviews, in which respondents are asked about their PID. In the case of ANES, a typical question is as follows: “Generally speaking, do you consider yourself a Democrat, Republican, Independent, or what?” Individual characteristics are observed, including sex, age, education and race (e.g., White, African, others).

There would be two strategies to estimate long-spanned PID dynamics. One is to conduct a one-shot estimation with the whole data. The other is to separate the whole periods into subperiods, and perform estimation multiple times with sub-period data. We take the latter strategy because it allows for finding more robust individual characteristics affecting PID mobility throughout the sample periods. Specifically, the entire three decade period is divided into three subperiods (the first period: 1972 to 1982, the second period: 1982 to 1992, the third period: 1992 to 2002) such that we can reasonably assume that individual PID evolves with a stable Markov dynamics in a decade. The sample sizes are as follows: 11,672 for the first subperiod, 12,196 for the second and 10,364 for the third. The initial distribution $\mathbf{p}(0)$ is constructed by PID observations in the initial year of each individual decade.

3.2 Bayesian Estimates

We estimate the dynamic properties of PID in the three decade. Following the standard practice in the Bayesian microeconometrics literature (see e.g., Rossi et al., 2005), we set prior parameters as: $v_0 = 5$, $\mathbf{V}_0 = v_0 \mathbb{I}_5$. The MCMC sampling is then conducted with four chains, each of which contains 20,000 iterations with the burn-in period of 10,000. The samples are saved after every two iterations, i.e., the effective iterations are 5,000 after the burn-in period. The convergence is confirmed with the diagnostic indicator \hat{R} taking the value of one for all parameters.

The explanatory variables include age, sex, education and race. The sex dummy takes “1” for females and “0” for males. The age variable is coded as “1” for the age group 17 -24 years old, “2” for 25 - 34, “3” for 35-44, “4” for 45-54, “5” for 55-64, “6” for 65-74 and “7” for 75-99 and over. The education dummy measures respondents’ completed education, coded as “1” for a post-college degree and “0” for less than high school education. The White dummy is coded “1” for White people and “0” for other races. The African dummy is coded as “1” for African Americans and “0” for other races.

Table 6 shows the results of the MCMC estimation. As the previous period’s partisan position is the benchmark, a coefficient measures one’s mobility to move to another partisan position, compared with retaining the benchmark position. For instance, if a variable of $\beta_{2|1}$ takes a negative coefficient, then the higher value of the variable makes an individual who was Democrat in the previous period less likely to be Independent (rather than Democrat) in the next period. Thus, a negative coefficient contributes to stable individual partisanship between subsequent periods, whereas a positive coefficient does the same to unstable partisanship. We interpret a coefficient estimate as statistically significant if its 95 % credible interval does not include zero.

Although the results in table 6 contain no parameters that are significant consistently over the three

⁶Wave 1 occurs in 1972, wave 2 is in 1974, and subsequent waves occur every two years.

decades studied, we find interesting propensities of individual partisanship. The column block for $\beta_{2|1}$ indicates that the stability of Democratic partisanship tends to be enhanced by age and education as well as in the group of African Americans. The column block for $\beta_{2|3}$ suggests that the stability of Republican partisanship tends to be enhanced in the group of White Americans, and reduced by education and in the group of African Americans. The column blocks for $\beta_{1|2}$ and $\beta_{3|2}$ reveal the dynamic propensity for the change in Independent partisanship. First, aging tends to reduce the stability of Independent partisanship. Specifically, older individuals who held an Independent position are more likely to move to Democratic or Republican positions in the next period. Second, focusing on the block for $\beta_{3|2}$, education is shown to influence individuals who were Independent in the previous period to move to the Republican position in the subsequent period. It also implies that African Americans with Independent partisanship tend to retain their partisanship, whereas White Americans with the same tend to move to the Republican position in the subsequent period. In short, the results support empirical evidence that age, education and race greatly affect the partisanship dynamics of American people.

Table 7 highlights our empirical evidence by illustrating differences, by race and education, in the dynamic patterns in the period of 1992 to 2002. The transitional matrices are resulted from the logistic functions (6), (7) and (11) and (12) with the mean values of significant estimators, controlling for other heterogeneities. As there are no significant parameters for White dummy variables, the partisanship transitional patterns of White Americans are indifferent from those of other Americans, otherwise with the same heterogeneities. As for African Americans, $\hat{\lambda}_{1|1}$ takes higher values, whereas $\hat{\lambda}_{2|2}$ and $\hat{\lambda}_{3|3}$ take lower values, which implies that African Americans have persistent partisanship for the Democratic Party but not for both the Independent partisanship and Republican Party. As for White and other Americans, $\hat{\lambda}_{1|1}$ and $\hat{\lambda}_{3|3}$ tend to take higher values than $\hat{\lambda}_{2|2}$, which implies that non-African Americans have persistent partisanship for the two parties but not for the Independent partisan. Nonetheless, for any race, high education enforces the persistence for Democratic partisanship, and reduces the same for Independent and Republican partisanship.

4 Conclusion

The dynamic property of party identification has been empirically investigated in political science and relevant economic studies. This study proposed a Bayesian approach for estimating the Markov dynamics of individual partisanship. The proposed method allows for investigation of partisan dynamics with long-spanning RCS data. The method is particularly useful when panel data are unavailable or inaccessible.

Using a data set covering 30 years, the study estimated the Markov dynamics for American individual partisanship based on the gradual transition theory in [Mebane and Wand \(1997\)](#): American individual partisanship gradually changes over time through the Independent status. The results find that age, education, and race significantly affect the transitional patterns of individual partisanship.

References

- Banks, Jeffrey and John Duggan**, “Probabilistic Voting in the Spatial Model of Elections: The Theory of Office-motivated Candidates,” in David Austen-Smith and John Duggan, eds., *Social Choice and Strategic Decisions*, Studies in Choice and Welfare, Springer Berlin Heidelberg, 2005, pp. 15–56. [2](#)
- Bartels, Brandon L, Janet M Box-Steffensmeier, Corwin D Smidt, and Renee M Smith**, “The dynamic properties of individual-level party identification in the United States,” *Electoral Studies*, 2011, *30* (1), 210–222. [2](#), [5](#)
- Campbell, James E., Mary Munro, John Alford, and Bruce A. Campbell**, “Partisanship and voting,” in Samuel Long, ed., *Research in Micropolitics*, JAI Press, 1986. [2](#)
- Clarke, Harold D. and Allan L. McCutcheon**, “The Dynamics of Party Identification Reconsidered,” *Public Opinion Quarterly*, 2009, *73* (4), 704–728. [2](#)
- Fiorina, Morris P. and Samuel J. Abrams**, “Political Polarization in the American Public,” *Annual Review of Political Science*, 2008, *11* (1), 563–588. [3](#)
- Gelman, Andrew and Donald B. Rubin**, “Inference from Iterative Simulation Using Multiple Sequences,” *Statistical Science*, 1992, *7* (4), 457–472. [7](#)
- Jackson, John E and Ken Kollman**, “Connecting Micro- and Macropartisanship,” *Political Analysis*, 2011, *19* (4), 503–518. [2](#)
- Mebane, Walter R. Jr and Jonathan Wand**, “Markov Chain Models for Rolling Cross-section Data: How Campaign Events and Political Awareness Affect vote Intentions and Partisanship in the United States and Canada,” in “in” the 1997 Annual Meeting of the Mid-west Political Science Association at Palmer House Hilton, Chicago, IL 1997. [2](#), [5](#), [9](#)
- Moffitt, Robert**, “Identification and estimation of dynamic models with a time series of repeated cross-sections,” *Journal of Econometrics*, 1993, *59* (1), 99–123. [3](#), [5](#)
- Neal, Radford M.**, “MCMC Using Hamiltonian Dynamics,” in “Handbook of Markov Chain Monte Carlo” Chapman & Hall/CRC Handbooks of Modern Statistical Methods, Chapman and Hall/CRC, 2011. [7](#)
- Pelzer, Ben, Rob Eisinga, and Philip Hans Franses**, “Inferring Transition Probabilities from Repeated Cross Sections,” *Political Analysis*, 2002, *10* (2), 113–133. [3](#)
- Persson, Torsten and Guido Tabellini**, *Political Economics: Explaining Economic Policy*, MIT Press, February 2002. [2](#)
- Rossi, Peter E., Greg M. Allenby, and Rob McCulloch**, *Bayesian Statistics and Marketing*, Wiley, 2005. [5](#), [8](#)

Table 1: RCS data structure

Individual index i	Period t	PID response y	Characteristics \mathbf{x}		
			Age	Sex	...
1	1	1	30'	Male	
2	1	3	20'	Female	
\vdots	\vdots	\vdots	\vdots	\vdots	\ddots
100	3	2	50'	Female	
\vdots	\vdots	\vdots	\vdots	\vdots	
I	τ	2	30'	Male	

Table 2: Transitional rates computed with a panel data set (Bartel *et al.*, 2011)

PID_{t-1}	PID_t		
	Democrat	Independent	Republican
Democrat	0.847	0.129	0.024
Independent	0.160	0.703	0.137
Republican	0.023	0.140	0.837

Entries are probabilities of being in a column state, conditional on being in a row state in the last period.

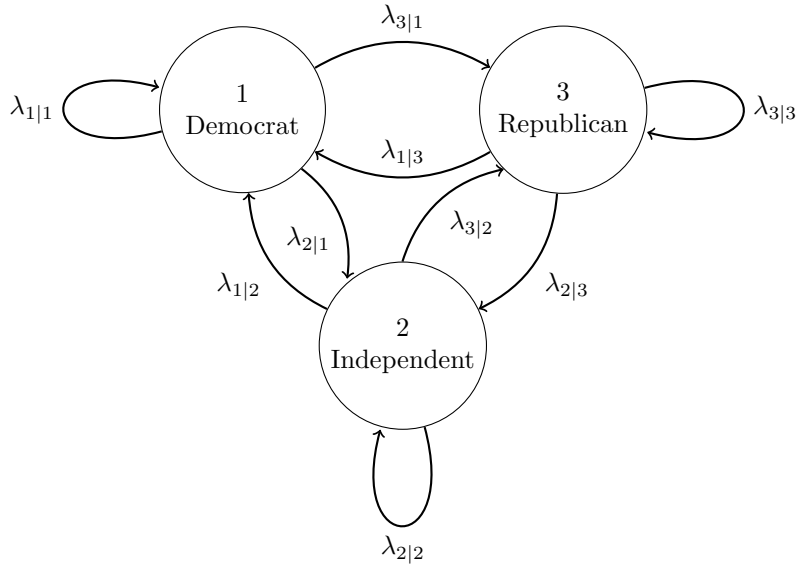


Figure 1: Three-state Markov chain

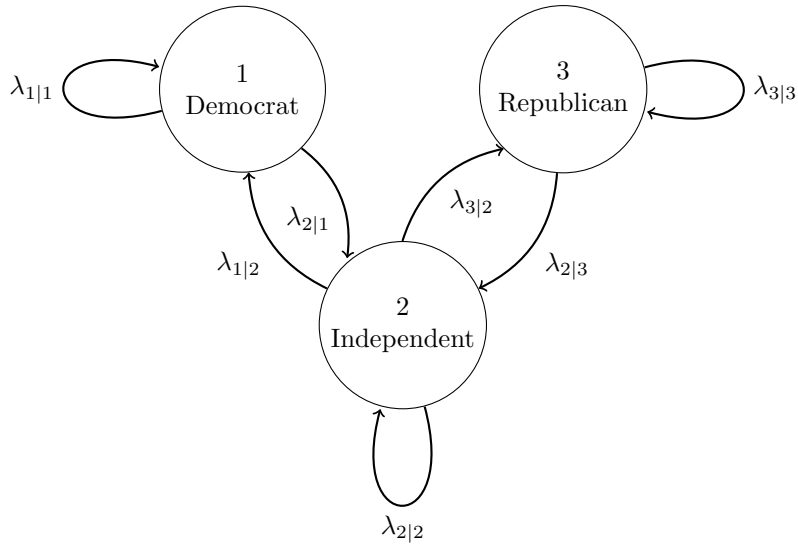


Figure 2: Gradually transitional Markov chain

Table 3: Types and its transitional matrices

			Type I			Type III		
			Characteristic A = 1			Characteristic A = 0		
			Characteristic B = 1			Characteristic B = 1		
$\lambda_{1 1}$	$\lambda_{2 1}$	$\lambda_{3 1}$	0.75	0.25	0	0.8	0.2	0
$\lambda_{1 2}$	$\lambda_{2 2}$	$\lambda_{3 2}$	0.1	0.8	0.1	0.3	0.4	0.3
$\lambda_{1 3}$	$\lambda_{2 3}$	$\lambda_{3 3}$	0	0.65	0.35	0	0.6	0.4
			Type II			Type IV		
			Characteristic A = 1			Characteristic A = 0		
			Characteristic B = 0			Characteristic B = 0		
$\lambda_{1 1}$	$\lambda_{2 1}$	$\lambda_{3 1}$	0.35	0.65	0	0.4	0.6	0
$\lambda_{1 2}$	$\lambda_{2 2}$	$\lambda_{3 2}$	0.1	0.8	0.1	0.3	0.4	0.3
$\lambda_{1 3}$	$\lambda_{2 3}$	$\lambda_{3 3}$	0	0.25	0.75	0	0.2	0.8

Table 4: Parameter posterior distributions in simulation

	β_{211}				β_{112}				β_{312}				β_{213}			
	Mean	St. Dev.	[2.5 %	Credibility Interval 50% 97.5 %]	Mean	St. Dev.	[2.5 %	Credibility Interval 50% 97.5 %]	Mean	St. Dev.	[2.5 %	Credibility Interval 50% 97.5 %]	Mean	St. Dev.	[2.5 %	Credibility Interval 50% 97.5 %]
Const.	0.49(*)	0.2	[0.09	0.49 0.87]	-0.7(*)	0.24	[-1.21	-0.7 -0.26]	-0.92(*)	0.24	[-1.37	-0.92 -0.43]	-2.9(*)	0.47	[-3.9	-2.88 -2.01]
Characteristic A	0.72(*)	0.22	[0.3	0.72 1.16]	-0.99(*)	0.29	[-1.54	-1 -0.42]	-1.23(*)	0.33	[-1.9	-1.22 -0.62]	1.58(*)	0.48	[0.69	1.56 2.59]
Characteristic B	-1.89(*)	0.28	[-2.49	-1.88 -1.38]	0.54	0.41	[-0.36	0.57 1.23]	0.59	0.34	[-0.05	0.59 1.28]	2.94(*)	0.49	[2.02	2.93 3.95]
$\mathbf{V}_{\beta,1,1}$					1.43	1.47	[0.42	1.12 4.32]	1.43	1.47	[0.42	1.12 4.32]				
$\mathbf{V}_{\beta,1,2}$					0.46	0.92	[-0.73	0.33 2.38]	0.46	0.92	[-0.73	0.33 2.38]				
$\mathbf{V}_{\beta,2,1}$					0.46	0.92	[-0.73	0.33 2.38]	0.46	0.92	[-0.73	0.33 2.38]				
$\mathbf{V}_{\beta,2,2}$					1.61	1.24	[0.48	1.26 4.86]	1.61	1.24	[0.48	1.26 4.86]				

¹ $\mathbf{V}_{\beta,k,\ell}$ is the k -th row and ℓ -th column of the covariance matrix \mathbf{V}_{β} .

² (*) indicates that the 95 % confidence interval excludes zero.

Table 5: Estimates and true values for transitional matrices

			Estimated Values			True Values		
Type I			(Characteristic A=1, Characteristic B=1)					
$\lambda_{1 1}$	$\lambda_{2 1}$	$\lambda_{3 1}$	0.66	0.34	0	0.75	0.25	0
$\lambda_{1 2}$	$\lambda_{2 2}$	$\lambda_{3 2}$	0.14	0.77	0.09	0.1	0.8	0.1
$\lambda_{1 3}$	$\lambda_{2 3}$	$\lambda_{3 3}$	0	0.83	0.17	0	0.65	0.35
Type II			(Characteristic A=1, Characteristic B=0)					
$\lambda_{1 1}$	$\lambda_{2 1}$	$\lambda_{3 1}$	0.23	0.77	0	0.35	0.65	0
$\lambda_{1 2}$	$\lambda_{2 2}$	$\lambda_{3 2}$	0.14	0.77	0.09	0.1	0.8	0.1
$\lambda_{1 3}$	$\lambda_{2 3}$	$\lambda_{3 3}$	0	0.21	0.79	0	0.25	0.75
Type III			(Characteristic A=0, Characteristic B=1)					
$\lambda_{1 1}$	$\lambda_{2 1}$	$\lambda_{3 1}$	0.80	0.20	0	0.8	0.2	0
$\lambda_{1 2}$	$\lambda_{2 2}$	$\lambda_{3 2}$	0.26	0.53	0.21	0.3	0.4	0.3
$\lambda_{1 3}$	$\lambda_{2 3}$	$\lambda_{3 3}$	0	0.51	0.49	0	0.6	0.4
Type IV			(Characteristic A=0, Characteristic B=0)					
$\lambda_{1 1}$	$\lambda_{2 1}$	$\lambda_{3 1}$	0.38	0.62	0	0.4	0.6	0
$\lambda_{1 2}$	$\lambda_{2 2}$	$\lambda_{3 2}$	0.26	0.53	0.21	0.3	0.4	0.3
$\lambda_{1 3}$	$\lambda_{2 3}$	$\lambda_{3 3}$	0	0.05	0.95	0	0.2	0.8

Table 6: Parameter posterior distributions

	β_{211}			β_{112}			β_{312}			β_{213}		
	Mean	St. Dev.	Credibility Interval [2.5 % 50 % 97.5 %]	Mean	St. Dev.	Credibility Interval [2.5 % 50 % 97.5 %]	Mean	St. Dev.	Credibility Interval [2.5 % 50 % 97.5 %]	Mean	St. Dev.	Credibility Interval [2.5 % 50 % 97.5 %]
1970												
Constant	-1.24(*)	0.8	[-2.99 -1.14 -0.10]	-1.8(*)	0.83	[-3.43 -1.8 -0.24]	-7.95(*)	1.81	[-12.00 -7.79 -4.85]	-0.97(*)	0.37	[-1.72 -0.96 -0.30]
Age	-0.18	0.09	[-0.34 -0.19 0.00]	1.07(*)	0.33	[0.61 1.03 1.87]	1.79(*)	0.43	[1.12 1.73 2.81]	-0.03	0.06	[-0.15 -0.03 0.10]
Sex	-0.34	0.23	[-0.75 -0.32 0.02]	0.4	0.48	[-0.54 0.4 1.36]	-0.03	0.56	[-1.11 -0.03 1.11]	-0.11	0.21	[-0.51 -0.12 0.32]
Education	-0.84(*)	0.25	[-1.30 -0.85 -0.33]	1.16	0.81	[-0.16 1.07 3.02]	6.1(*)	1.38	[3.68 6 9.07]	0.38	0.3	[-0.21 0.38 0.97]
White	0.89	0.64	[-0.02 0.82 2.18]	0.12	0.75	[-1.36 0.11 1.65]	0.19	1.21	[-1.99 0.11 2.88]	-1.56(*)	0.39	[-2.31 -1.57 -0.76]
African	-3.98(*)	2.76	[-11.00 -3.15 -0.73]	0.49	0.83	[-0.96 0.42 2.29]	-4.49(*)	2	[-9.04 -4.3 -1.16]	0.88(*)	0.33	[0.27 0.87 1.55]
$V_{\beta,1,1}$				1.77	1.36	[0.51 1.42 5.04]	1.77	1.36	[0.51 1.42 5.04]			
$V_{\beta,1,2}$				2.96	3.12	[-0.85 2.26 10.42]	2.96	3.12	[-0.85 2.26 10.42]			
$V_{\beta,2,1}$				2.96	3.12	[-0.85 2.26 10.42]	2.96	3.12	[-0.85 2.26 10.42]			
$V_{\beta,2,2}$				17.55	12.69	[4.81 14.17 50.18]	17.55	12.69	[4.81 14.17 50.18]			
1980												
Constant	-0.55	0.42	[-1.48 -0.51 0.16]	0.65	1.07	[-1.15 0.53 3.15]	-1.25	1.07	[-3.71 -1.14 0.55]	-2.16(*)	0.51	[-3.21 -2.15 -1.20]
Age	-0.24(*)	0.05	[-0.34 -0.24 -0.14]	0.67(*)	0.34	[0.20 0.62 1.50]	0.52(*)	0.35	[0.03 0.46 1.41]	-0.02	0.05	[-0.12 -0.02 0.09]
Sex	-0.37(*)	0.14	[-0.68 -0.36 -0.11]	-0.48	0.67	[-1.76 -0.5 0.87]	-0.5	0.61	[-1.69 -0.51 0.76]	0.3	0.19	[-0.07 0.3 0.67]
Education	-1.4(*)	0.2	[-1.77 -1.42 -0.96]	-1.84	1.66	[-5.28 -1.8 1.26]	3.77(*)	1.62	[1.28 3.52 7.63]	0.73(*)	0.3	[0.22 0.71 1.40]
White	0.61(*)	0.31	[0.06 0.58 1.27]	-0.39	0.94	[-2.54 -0.31 1.23]	0.92	0.97	[-0.65 0.81 3.19]	-1.08(*)	0.44	[-1.89 -1.09 -0.17]
African	-0.75	0.41	[-1.47 -0.78 0.14]	3.32(*)	2.07	[0.17 3 8.14]	-3.52(*)	2.08	[-8.45 -3.21 -0.30]	1.13(*)	0.45	[0.31 1.11 2.07]
$V_{\beta,1,1}$				3.78	4.72	[0.57 2.45 15.14]	3.78	4.72	[0.57 2.45 15.14]			
$V_{\beta,1,2}$				-2.07	3.08	[-9.18 -1.44 1.56]	-2.07	3.08	[-9.18 -1.44 1.56]			
$V_{\beta,2,1}$				-2.07	3.08	[-9.18 -1.44 1.56]	-2.07	3.08	[-9.18 -1.44 1.56]			
$V_{\beta,2,2}$				5.47	5.52	[0.87 3.83 20.00]	5.47	5.52	[0.87 3.83 20.00]			
1990												
Constant	-0.65	0.68	[-2.20 -0.56 0.26]	-0.23	2.14	[-3.23 -0.71 5.19]	-4.6(*)	2.11	[-8.51 -4.74 -0.05]	-3.29(*)	0.56	[-4.49 -3.23 -2.36]
Age	-0.25(*)	0.08	[-0.40 -0.25 -0.07]	1.18	0.7	[-0.01 1.09 2.81]	1.35(*)	0.72	[0.25 1.25 3.06]	0.06	0.07	[-0.07 0.06 0.21]
Sex	-0.3	0.22	[-0.72 -0.3 0.14]	-0.14	0.97	[-1.97 -0.17 1.91]	-1.26	1.11	[-3.57 -1.2 0.78]	0.44	0.24	[-0.04 0.44 0.89]
Education	-1.38(*)	0.39	[-1.98 -1.42 -0.40]	2.13	2.55	[-3.92 2.52 6.36]	6.91(*)	2.73	[1.31 7.06 12.20]	0.64(*)	0.29	[0.13 0.62 1.27]
White	0.55	0.42	[-0.07 0.52 1.29]	0.6	1.49	[-2.42 0.6 3.65]	2.95	1.84	[-0.08 2.77 7.12]	-0.29	0.67	[-1.27 -0.27 0.74]
African	-4.26(*)	2.87	[-11.35 -3.59 -0.84]	-0.41	2.15	[-3.78 -0.76 4.86]	-4.75(*)	2.6	[-10.54 -4.55 -0.25]	2.34(*)	0.48	[1.60 2.28 3.41]
$V_{\beta,1,1}$				3.85	5.48	[0.67 2.53 14.43]	3.85	5.48	[0.67 2.53 14.43]			
$V_{\beta,1,2}$				4.77	6.55	[-3.49 3.57 20.70]	4.77	6.55	[-3.49 3.57 20.70]			
$V_{\beta,2,1}$				4.77	6.55	[-3.49 3.57 20.70]	4.77	6.55	[-3.49 3.57 20.70]			
$V_{\beta,2,2}$				16.57	14.93	[1.28 12.9 55.10]	16.57	14.93	[1.28 12.9 55.10]			

¹ $V_{\beta,k,\ell}$ is the k -th row and ℓ -th column of the covariance matrix V_{β} .

² (*) indicates that the 95 % confidence interval excludes zero.

Table 7: Estimated transitional matrices for the period of 1992-2002

			African Americans with below high-school education (education dummy = 0, African dummy = 1)			White/ Other Americans with below high-school education (education dummy = 0, African dummy = 0)		
$\hat{\lambda}_{1 1}$	$\hat{\lambda}_{2 1}$	$\hat{\lambda}_{3 1}$	0.993	0.007	0	0.679	0.321	0
$\hat{\lambda}_{1 2}$	$\hat{\lambda}_{2 2}$	$\hat{\lambda}_{3 2}$	0.499	0.499	0.002	0.388	0.388	0.224
$\hat{\lambda}_{1 3}$	$\hat{\lambda}_{2 3}$	$\hat{\lambda}_{3 3}$	0	0.316	0.684	0	0.043	0.957
			African Americans with above college education (education dummy = 1, African dummy = 1)			White/ Other Americans with above college education (education dummy = 1, African dummy = 0)		
$\hat{\lambda}_{1 1}$	$\hat{\lambda}_{2 1}$	$\hat{\lambda}_{3 1}$	0.998	0.002	0	0.894	0.106	0
$\hat{\lambda}_{1 2}$	$\hat{\lambda}_{2 2}$	$\hat{\lambda}_{3 2}$	0.143	0.143	0.714	0.002	0.002	0.996
$\hat{\lambda}_{1 3}$	$\hat{\lambda}_{2 3}$	$\hat{\lambda}_{3 3}$	0	0.468	0.532	0	0.078	0.922

¹ Significant parameters are set to the mean values, otherwise zero.

² The value of the control variable, age is set to three.

Appendix A Latent variable interpretation for multinomial logistic function

This appendix details how the parameters of interest $\beta_{\ell|k}$ in our econometric model relate to the utility maximization of individuals. We assume that individual i seeks to maximize his/ her utility by choosing PID $\ell \in \Omega$, given the last period's PID choice $k \in \Omega$. The maximized utility is then given by:

$$U_{i\ell} := \mathbf{x}'_i \beta_{\ell|k} + \epsilon_{i\ell}$$

where $\epsilon_{i\ell}$ is an error term, and follows an i.i.d. type I extreme value distribution. The response probability that individual i will choose PID $\ell \in \Omega$ is then given by:

$$\begin{aligned} \lambda_{\ell|k}^i &= \text{Prob}(U_{i\ell} \geq U_{ik}) \quad \text{for all } k \neq \ell \\ &= \text{Prob}(U_{i\ell} - U_{ik} \geq 0) \quad \text{for all } k \neq \ell \\ &= \text{Prob}(\epsilon_{ik} - \epsilon_{i\ell} \leq \mathbf{x}'_i \beta_{\ell|k} - \mathbf{x}'_i \beta_{k|k}) \quad \text{for all } k \neq \ell \\ &= \begin{cases} \frac{\exp(\mathbf{x}'_i \beta_{\ell|k})}{\sum_{m=1}^3 \exp(\mathbf{x}'_i \beta_{m|k})}, & \text{for } \ell \neq k \\ \frac{1}{\sum_{m=1}^3 \exp(\mathbf{x}'_i \beta_{m|k})}, & \text{for } \ell = k \end{cases} \end{aligned}$$