19

Modeling Preference Changes via a Hidden Markov Item Response Theory Model

Jong Hee Park

CONTENTS

19.1 Introduction

Over the past two decades, political scientists have made great advances in the empirical estimation of ideal points (Bafumi et al., 2005; Bailey and Chang, 2001; Clinton et al., 2000; Heckman and Snyder, 1997; Jackman, 2001; Londregan, 2000; Martin and Quinn, 2002; Poole and Rosenthal, 1997). An ideal point, or preference, is a foundational theoretical concept for explaining the choices a political actor makes. For example, in simple unidimensional spatial models of voting, a legislator's vote choice is modeled as a rational decision based on a (Euclidean geometric) calculation of differences in utility values between the legislator's ideal point, a proposed bill, and the status quo.

Although an ideal point is often assumed to be static for theoretical convenience, dynamics in ideal points pose an important theoretical and empirical puzzle to researchers. For example, examining the judicial opinion writing of 16 US Supreme Court justices, Epstein et al. (1998) conclude that there is enough evidence to invalidate the assumption of preference stability over time.* They also go on to claim that any inference about a justice's "revealed preference" that is based on the stable preference assumption can be misleading if the justice actually underwent several preference changes over a lifetime. The development of statistical methods for dynamic ideal point estimation has been limited to a few published works (Martin and Quinn, 2002; McCarty et al., 1997). Also, the existing methods

* The study of linkages between judges' opinions and their ideological leanings has become an important area of research in the last two decades, as political scientists have rejected "apolitical" legal understandings of judicial opinions in favor of attitudinal and rational models that introduce "political" factors into the decision-making process (Epstein and Knight, 1998; Segal and Spaeth, 1993).

for dynamic ideal point estimation fail to distinguish fundamental changes from random drifts. In this paper I propose a method to detect *sharp, discontinuous changes in ideal points*.

The approach I take in this paper is to combine Chib's (1998) hidden Markov model (HMM) with the two-parameter item response theory (IRT) model. In this model, the dynamics in ideal points are modeled as agent-specific hidden regime changes. I demonstrate the utility of the hidden Markov IRT model by analyzing changes in ideal points among the 43 US Supreme Court justices serving between 1937 and 2006, and conclude that the model provides an effective benchmark for making probabilistic inferences about the timing of preference changes.

19.2 Dynamic Ideal Point Estimation

Assuming a quadratic utility loss function, the utility of voting for item i by legislator j at time t is

$$U_{jt}(Y_i) = -(\theta_{jt} - Y_i)^2 + \delta_{ijt}^{(Y)},$$

where θ_{jt} is legislator j's ideal point at t, Y_i is the location of Yay, and $\delta^{(Y)}$ is a stochastic error drawn from a Gaussian distribution. For simple notation, I assume that θ_{jt} and Y_i are scalar, which means that the underlying political space is one-dimensional. The utility of voting against item i is defined similarly:

$$U_{jt}(N_i) = -(\theta_{jt} - N_i)^2 + \delta_{ijt}^{(N)}.$$

In this random utility model, a legislator votes for a bill i when $U_{jt}(Y_i) - U_{jt}(N_i) > 0$. If the utility difference between two vote choices is treated as a latent variable, the process can be modeled as a Bernoulli trial in which the probability of a yes vote is a function of a legislator's ideal point and the proposed bill's location:

$$y_{ijt} = \begin{cases} 1, & \text{if } z_{ijt} = U_{jt}(Y_i) - U_{jt}(N_i) > 0, \\ 0, & \text{if } z_{ijt} = U_{jt}(Y_i) - U_{jt}(N_i) \leq 0. \end{cases}$$

Then, as shown by Jackman (2001), some simple algebra shows the connection between the random utility voting model and the two-parameter IRT model:

$$\begin{aligned} z_{ijt} = U_{jt}(Y_i) - U_{jt}(N_i) &= -(\theta_{jt} - Y_i)^2 + \delta_{ijt}^{(Y)} + (\theta_{jt} - N_i)^2 - \delta_{ijt}^{(N)} \\ &= -2(N_i - Y_i)\theta_{jt} - (Y_i^2 - N_i^2) + \delta_{ijt}^{(Y)} - \delta_{ijt}^{(N)} \\ &= \beta_i\theta_{jt} - \alpha_i + \varepsilon_{ijt}. \end{aligned}$$

Note that the t subscript in ideal points is carried through the equation to denote the dynamics in ideal points.

If a political actor only makes a few decisions or is active for a short period of time, ignoring idea point temporal dynamics is unlikely to pose a problem. However, for someone such as a legislator who serves multiple terms in office sessions, the conventional IRT model with constant ideal points very likely fails to capture any political evolution. As the legislator ages, exposure to exogenous shocks in the form of economic shifts, social upheavals, and new political environments is likely to affect voting decisions. It would be unrealistic to attribute all time-varying patterns in voting behavior to bill characteristics (α and β).*

Since the constant IRT model itself is highly parameterized with $2I + J$ parameters, where I is the number of items and J is the number of legislators, letting ideal points (θ_{jt}) vary over time is not a trivial modification. Two methods have been proposed so far. The first method, which does not rely on the IRT framework, is to specify ideal points as a polynomial function of time (McCarty et al., 1997). The other is to model the transition of ideal points as a first-order Markov process while the observed voting data are generated from the IRT model (Martin and Quinn, 2002). One major difference between two methods is the source of the dyanmics. McCarty et al.'s (1997) method assumes that the effect of time on ideal points is deterministic. By contrast, Martin and Quinn's dynamic ideal point method decomposes the source of changes in ideal points into a deterministic part and a stochastic part, and estimates the variance of legislator-specific transitions. These two methods are successfully applied to developing dynamic measurements of ideal points in US legislators and US Supreme Court justices with DW-NOMINATE (McCarty et al., 1997) and the Martin–Quinn score (Martin and Quinn, 2002), respectively.

However, while both methods are effective in uncovering transitions in ideal points, neither is specifically designed to detect the timing of changes in ideal points. In other words, the existing dynamic ideal point estimation methods are not optimal for modeling sharp, discontinuous changes in ideal points. This is an important issue since theoretical discussions on changes in ideal points pit continuous transitions of ideal points against discontinuous transitions. To put it differently, researchers who are more interested in the existence and timing of ideal point shifts rather than with smooth evolutions of ideal points over time would not find the existing methods helpful. This is why I have introduced a dynamic IRT model specifically designed to capture sharp, discontinuous ideal point shifts.

19.3 Hidden Markov Item Response Theory Model

The approach I take combines a HMM with the standard two parameter IRT model. Specifically, I use Chib's (1998) model to capture hidden regime changes in a legislator's ideal point. Note that in Chib's model the regime transition is constrained so that a Markov chain only moves forward to the terminal state. This constraint generates a nonergodic Markov chain, which turns out to be computationally efficient and as flexible as HMMs with ergodic Markov chains.†

* Note that the variance parameter in the IRT model is not identified as in the binary response models.

† Let $p_{ij} = \Pr(s_t = j \mid s_{t-1} = i)$ be the probability of moving to state j from state i at time t when the state at $t-1$ is i. Then, the transition matrix of Chib (1998) is

$$\mathbf{P} = \begin{pmatrix} p_{11} & p_{12} & 0 & \cdots & 0 \\ 0 & p_{22} & p_{23} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & p_{M-1,M-1} & p_{M-1,M} \\ 0 & 0 & & 0 & 1 \end{pmatrix}.$$

Let s_{jt} be an indicator of hidden regimes for legislator j's ideal point at t, and P_j be a transition matrix for the hidden regimes. Due to the nonergodic constraint, it is trivial to compute the initial probability: $\pi_0 = (1, \ldots, 0)$. The latent propensity of voting for an item i can be expressed as a function of item characteristics and ideal points, which are subject to agent-specific regime changes:

$$z_{ijk} = \beta_i \theta_{j,s_{jt}} - \alpha_i + \varepsilon_{ijt}, \quad \varepsilon_{ijt} \sim N(0,1), \tag{19.1}$$

$$s_{jt} \mid s_{j,t-1} \sim \text{Markov}(\pi_0, P_j) \tag{19.2}$$

The value of $\theta_{j,s_{jt}}$ can take M different values at each time point subject to the first-order Markov process. In other words, s_{jt} indicates the preference regime associated with a legislator's ideal point at t.

It should be stressed that the hidden Markov IRT model can be considered as a special type of the dynamic IRT model developed by Martin and Quinn (2002). While the dynamic IRT model assumes that ideal points change at each time point due to random shocks, the hidden Markov IRT model assumes that ideal points change only when the underlying regime changes. When there is no detected change point, $s_{jt} = 1$ for $t = 1, \ldots, T_j$, the hidden Markov IRT model reduces to the constant IRT model.

Albert (1992) and Johnson and Albert (1999) provide an efficient Gibbs sampling algorithm for the constant IRT model. Once hidden state variables s_{jt} are sampled, the rest of the sampling scheme is similar to the constant IRT model.

Normal distributions are used as prior distributions for ideal points and item parameters. For identification, I use the standard normal distribution as a prior distribution of ideal points:*

$$\begin{aligned} \boldsymbol{\lambda}_i &\sim N(\mu_0, \mathbf{V}_0), \\ \theta_{j,s_t} &\sim N(0,1), \\ p_{ii} &\sim \text{Beta}(a,b), \end{aligned}$$

where $\boldsymbol{\lambda}_i = (\alpha_i, \beta_i)'$.

The MCMC sampling algorithm for the hidden Markov IRT model consists of five steps, including two steps for augmented variables. We have

$$\begin{aligned} p(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\theta}, \mathbf{P} \mid \mathbf{y}) &= \int p(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\theta}, \mathbf{P}, \mathbf{s}, \mathbf{z} \mid \mathbf{y})\, d\mathbf{s}\, d\mathbf{z} \\ &= \int p(\boldsymbol{\alpha}, \boldsymbol{\beta} \mid \boldsymbol{\theta}, \mathbf{P}, \mathbf{s}, \mathbf{z}) p(\boldsymbol{\theta} \mid \mathbf{P}, \mathbf{s}, \mathbf{z}) p(\mathbf{P} \mid \mathbf{s}, \mathbf{z}) p(\mathbf{s} \mid \mathbf{z}, \mathbf{y}) p(\mathbf{z} \mid \mathbf{y})\, d\mathbf{s}\, d\mathbf{z}. \end{aligned}$$

Step 1. Simulation of latent utilities. Following Albert and Chib (1993), the latent variable (z_{ijk}) in Equation 19.1 is sampled from two truncated normal distributions, the support of which changes depending on realized binary outcomes:

$$z_{ijk} \sim \begin{cases} N_{(-\infty,0]}(\alpha_i + \beta_i \theta_{j,s_{jt}}, 1), & \text{if } y_{ijt} = 0, \\ N_{(0,\infty)}(\alpha_i + \beta_i \theta_{j,s_{jt}}, 1), & \text{if } y_{ijt} = 1. \end{cases}$$

* See Clinton et al. (2000) and Jackman (2001) for the identification of the IRT model in Bayesian estimation.

Step 2. Simulation of item parameters. A vectorized notation is used to explain the simulation of item parameters. Latent utilities are formed as a $J \times 1$ vector ($\mathbf{z}_{it}$), and ideal point estimates are transformed into a $J \times 2$ matrix $\mathbf{\Theta}_t = (\mathbf{1}, \boldsymbol{\theta}_{j,s_{jt}})$ where $\boldsymbol{\theta}_{j,s_{jt}}$ is a vector of ideal points for all legislators at time t: $\boldsymbol{\theta}_{j,s_{jt}} = (\theta_{1,s_{1,t}}, \ldots, \theta_{J,s_{J,t}})'$. Finally, item parameters are stacked as a 2×1 matrix $\boldsymbol{\lambda}_i = (\alpha_i, \beta_i)'$. Then we have a multivariate linear regression model

$$\mathbf{z}_{it} = \mathbf{\Theta}_t \boldsymbol{\lambda}_i + \boldsymbol{\varepsilon}_{it}.$$

Note that $\mathbf{\Theta}_t$ serves as a design matrix and $\boldsymbol{\lambda}_i$ serves as a parameter vector at this sampling step. From this,

$$\boldsymbol{\lambda}_i \mid \boldsymbol{\theta}, \mathbf{P}, \mathbf{z}, \mathbf{y} \sim N(\mu_{\boldsymbol{\lambda}}, \mathbf{V}_{\boldsymbol{\lambda}}),$$

$$\mathbf{V}_{\boldsymbol{\lambda}} = \left(\sum_{t=1}^{T} \mathbf{\Theta}_t' \mathbf{\Theta}_t + \mathbf{V}_0^{-1} \right)^{-1},$$

$$\mu_{\boldsymbol{\lambda}} = \mathbf{S}_{\boldsymbol{\lambda}} \left(\sum_{t=1}^{T} \mathbf{\Theta}_t' \mathbf{z}_{it} + \mathbf{V}_0^{-1} \mu_0 \right).$$

Step 3. Simulation of a latent state vector. For the simulation of ideal points, I transform Equation 19.1 into a multivariate time series model by subtracting the difficulty parameter $\boldsymbol{\alpha}_t^j$ from latent utilities ($\mathbf{z}_{jt}^* = \mathbf{z}_{jt} - \boldsymbol{\alpha}_t^j$) and stacking them as an $I_t^j \times 1$ vector. I_t^j indicates the number of items considered by legislator j at time t and varies across legislators. $\boldsymbol{\alpha}_t^j$ indicates difficulty parameters for all items considered by legislator j at time t. The dimension of $\boldsymbol{\alpha}_t^j$ also changes by legislators and time. Similarly, let $\boldsymbol{\beta}_t^j$ denote discrimination parameters for all items considered by legislator j at time t. The new equation can take the form of a linear regression model with $\boldsymbol{\beta}_t^j$ as a $I_t^j \times 1$ design matrix and θ_{j,s_t} as a 1×1 parameter vector as follows:

$$\mathbf{z}_{jt}^* = \boldsymbol{\beta}_t^j \theta_{j,s_t} + \boldsymbol{\varepsilon}_{jt}. \tag{19.3}$$

Sampling a latent state vector for each legislator is done using Chib's (1998) recursive sampling algorithm. The algorithm is identical for all legislators and needs to be repeated J times. Thus, I drop subscript j for notational simplicity. $\boldsymbol{\beta}, \mathbf{z}^*, \boldsymbol{\theta}$ and P should be read as $\boldsymbol{\beta}^j, \mathbf{z}_j^*, \boldsymbol{\theta}_j$ and P_j in the following. Note that I suppress time subscripts of $\boldsymbol{\beta}, \mathbf{z}^*, \boldsymbol{\theta}$ to denote them as matrices containing all observations. The joint sampling of latent states can be decomposed as follows:

$$\begin{aligned} p(s_1, \ldots, s_T \mid \boldsymbol{\beta}, \mathbf{z}^*, \boldsymbol{\theta}, P) &= p(s_T \mid \boldsymbol{\beta}, \mathbf{z}^*, \boldsymbol{\theta}, P) p(s_{T-1}, s_{T-2}, \ldots, s_1 \mid s_T, \boldsymbol{\beta}, \mathbf{z}^*, \boldsymbol{\theta}, P) \\ &= p(s_T \mid \boldsymbol{\beta}, \mathbf{z}^*, \boldsymbol{\theta}, P) \ldots p(s_t \mid S^{t+1}, \boldsymbol{\beta}, \mathbf{z}^*, \boldsymbol{\theta}, P) \ldots \\ &\quad p(s_1 \mid \boldsymbol{\beta}, \mathbf{z}^*, S^2, \boldsymbol{\theta}, P), \end{aligned} \tag{19.4}$$

where $\mathbf{S}^{t+1}$ indicates the history of the state from $t+1$ to T. Using Bayes' theorem, a typical form of Equation 19.4 can be decomposed as follows:

$$p(s_t \mid S^{t+1}, \boldsymbol{\beta}, \mathbf{z}^*, \boldsymbol{\theta}, P) \propto p(s_{t+1} \mid s_t, P) p(s_t \mid \boldsymbol{\beta}, \mathbf{z}_t^*, \boldsymbol{\theta}, P).$$

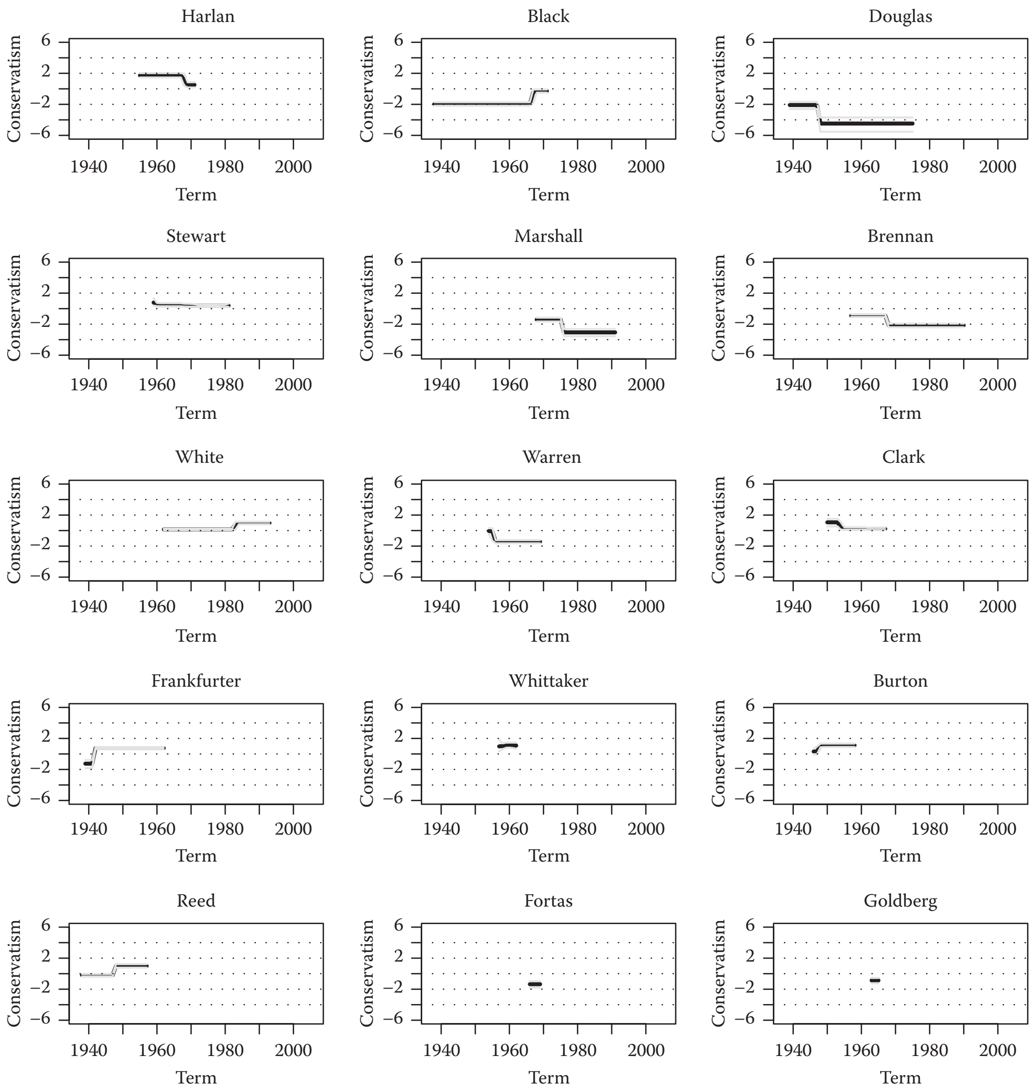

FIGURE 19.1
Estimated ideal points with one break. The chain was run for 20,000 draws after throwing out the first 10,000. Every 10th draw was stored for the analysis. Thick lines indicate posterior means and light lines are 95% Bayesian credible intervals.

The first part of the right-hand side is a transition probability from t to $t+1$, which is obtained from a transition matrix (P). The second part of the right-hand side should be obtained via recursive calculation. Let $\mathbf{Z}_t^*$ denote all $\mathbf{z}^*$ up to t. Then

$$\begin{aligned} p(s_t \mid \mathbf{Z}_{t-1}^*, \boldsymbol{\beta}, \boldsymbol{\theta}, P) &= \int_{\mathbf{S}} p(s_t \mid s_{t-1}) p(s_{t-1} \mid \mathbf{Z}_{t-1}^*, \boldsymbol{\beta}, \boldsymbol{\theta}, P) ds_t \\ &= \sum_{m=1}^{M} p(s_t \mid s_{t-1} = m) p(s_{t-1} = m \mid \mathbf{Z}_{t-1}^*, \boldsymbol{\beta}, \boldsymbol{\theta}, P), \end{aligned}$$

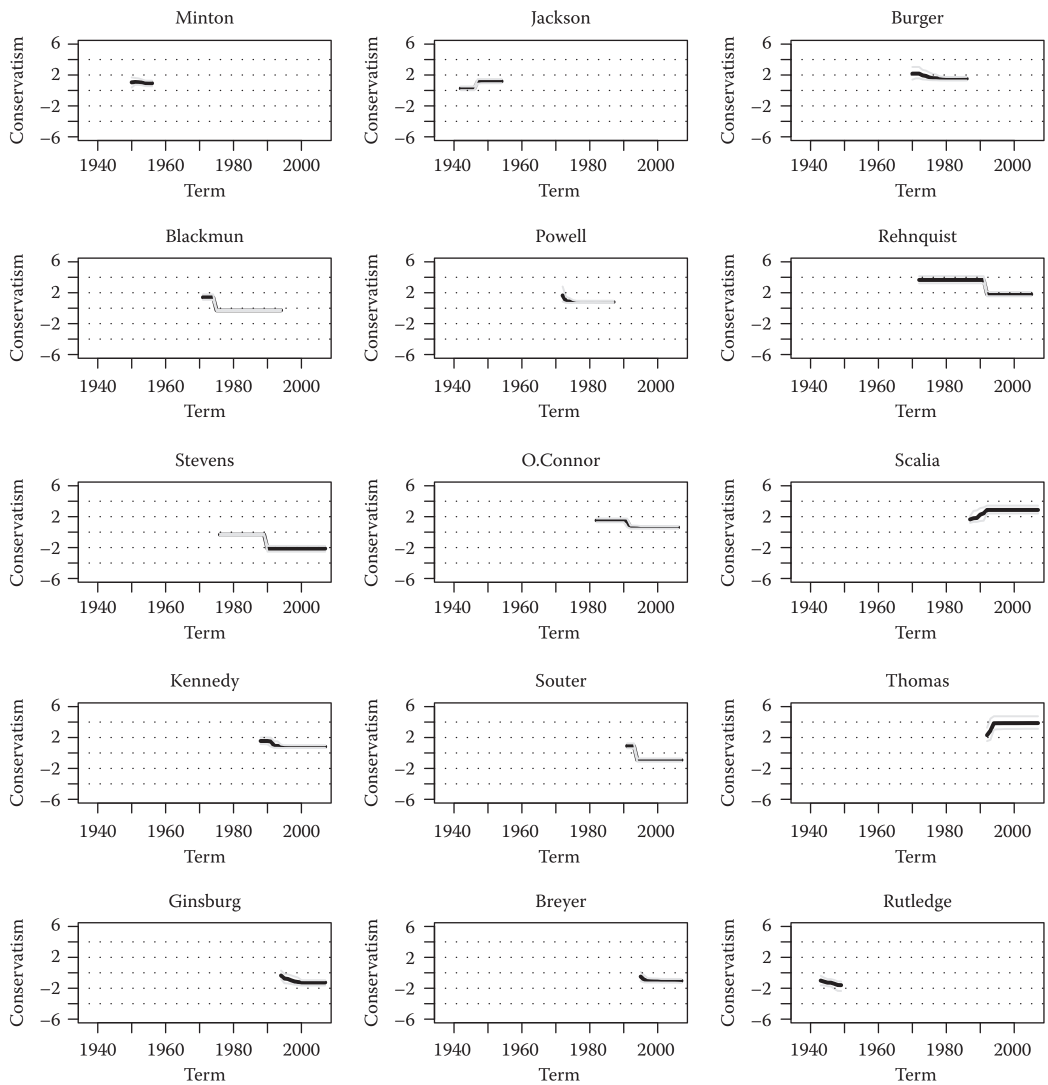

FIGURE 19.1
Continued.

$$p(s_t \mid \mathbf{Z}_t^*, \boldsymbol{\beta}, \boldsymbol{\theta}, P) = \frac{p(s_t \mid \mathbf{Z}_{t-1}^*, \boldsymbol{\beta}, \boldsymbol{\theta}, P)p(\mathbf{z}_t^* \mid \mathbf{Z}_{t-1}^*, \theta_{s_t})}{\sum_{m=1}^{M} p(s_t = m \mid \mathbf{Z}_{t-1}^*, \boldsymbol{\beta}, \boldsymbol{\theta}, P)p(\mathbf{z}_t^* \mid \mathbf{Z}_{t-1}^*, \theta_{s_t=m})}.$$

Step 4. Simulation of transition probabilities. Simulating transition probabilities given sampled state variables is a standard beta update from binary outcomes. For each legislator, let n_{ii} be the number of one-step transitions from state i to i, and n_{ij} be the number of one-step transitions from state i to j. Then for the posterior distributions of legislator-specific transition probabilities we have

$$\begin{aligned} p(p_{ii} \mid \mathbf{s}) &\propto p(\mathbf{s} \mid p_{ii})\text{Beta}(a, b) \\ &\propto p_{ii}^{n_{ii}}(1 - p_{ii})^1 p_{ii}^{a-1}(1 - p_{ii})^{b-1}, \\ p_{ii} \mid \mathbf{s} &\sim \text{Beta}(a + n_{ii}, b + 1). \end{aligned}$$

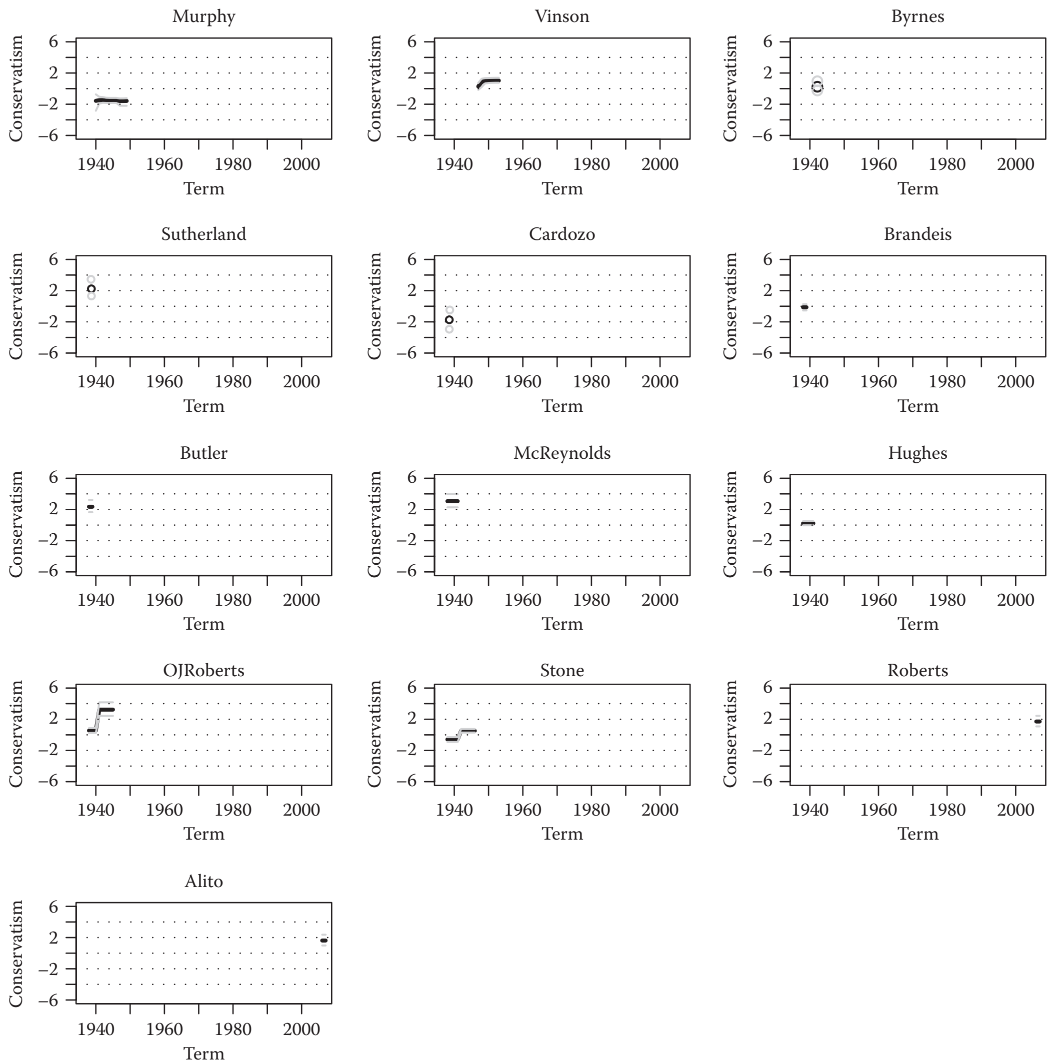

FIGURE 19.1
Continued.

Step 5. Simulation of ideal points. Ideal points are sampled using the transformation shown in Equation 19.3. Based on the sampled transition matrices, state variables, item characteristic parameters, and latent variables, θ_{j,s_t} is updated across all legislators by treating $\boldsymbol{\beta}^j$ as a design matrix and $\mathbf{z}^*_{j,t}$ as response variables. Let $\boldsymbol{\beta}_{j,m}$ and $\mathbf{z}^*_{j,m}$ denote sampled parameters for legislator j's mth state. Then

$$\theta_{j,m} \mid \mathbf{z}^*_{j,m} \sim N(\mu_\theta^{j,m}, \mathbf{V}_\theta^{j,m}),$$

$$\mathbf{V}_\theta^{j,m} = (\boldsymbol{\beta}'_{j,m}\boldsymbol{\beta}_{j,m} + 1)^{-1},$$

$$\mu_\theta^{j,m} = \mathbf{V}_\theta^{j,m}(\boldsymbol{\beta}'_{j,m}\mathbf{z}^*_{j,m}).$$

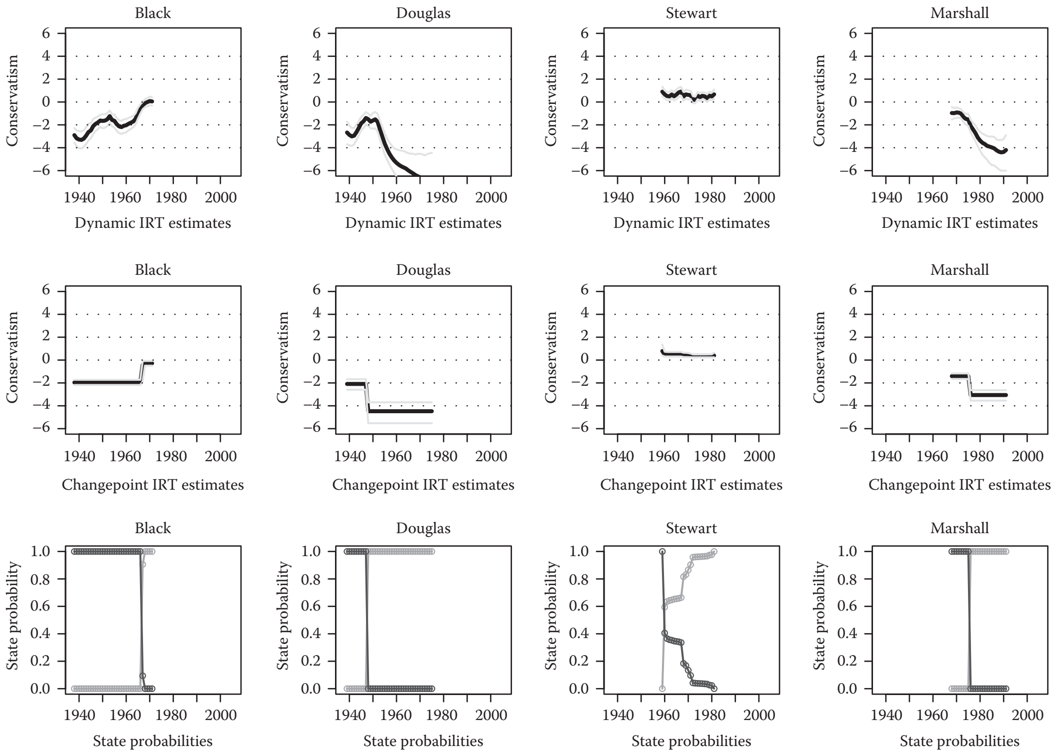

FIGURE 19.2
Comparison of ideal point estimates from the dynamic IRT model by Martin and Quinn (2002) and the hidden Markov IRT model: Black, Douglas, Stewart, and Marshall. The chain was run for 20,000 draws after throwing out the first 10,000. Thick lines on the plots in the top two rows indicate posterior means and light lines are 95% Bayesian credible intervals. The plots in the bottom row show posterior probabilities of being in state 1 (light lines) and in state 2 (dark lines).

19.4 Preference Changes in US Supreme Court Justices

Using the hidden Markov IRT model, I analyze ideal point changes of 43 U.S. Supreme Court justices who served between 1937 and 2006.* The 43 justices considered 4868 cases during the period, and on average each justice considers 113 cases throughout their terms on the bench.

In this analysis, I drop six justices with 2 years of service or less: Sutherland, Cardozo, Brandeis, Butler, Roberts, and Alito. Following Martin and Quinn (2007), I use informative priors for three justices, the liberal Hugo Black, the moderate Potter Stewart, and the conservative William Rehnquist, in order to interpret results in such a way that positive ideal point estimates indicate the conservative position and negative ideal point estimates indicate the liberal position:

$$\theta_{\text{Black}} \sim N(-2, 0.1),$$

* I thank Martin and Quinn for providing data. For details, see Martin and Quinn (2007).

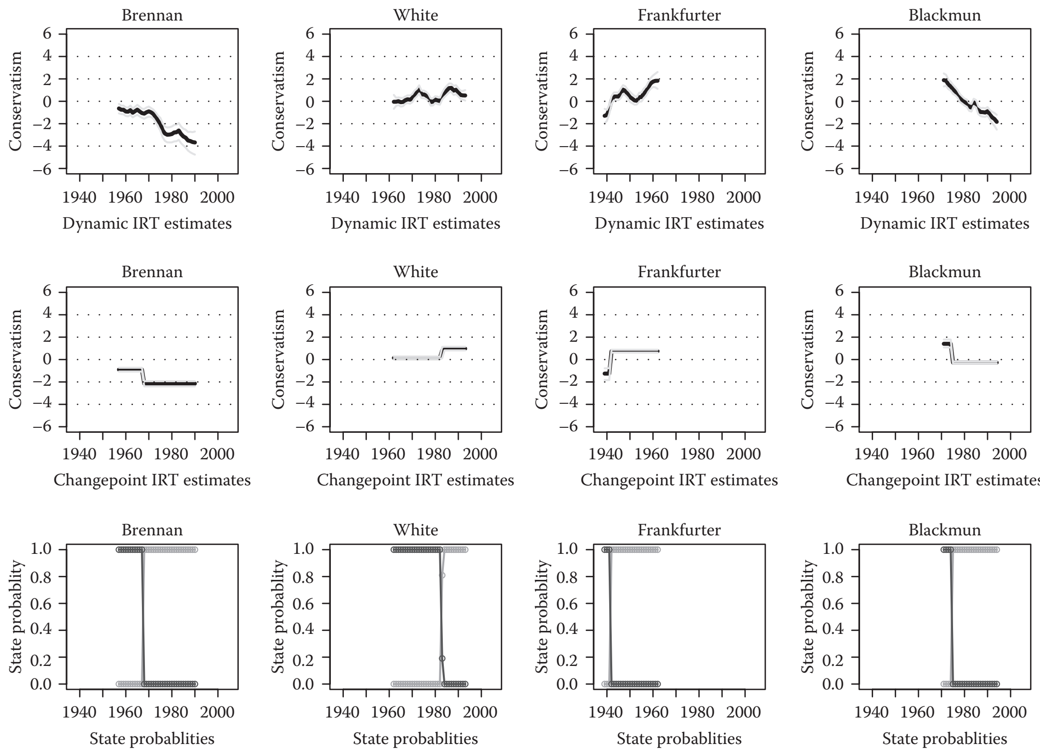

FIGURE 19.3
Comparison of ideal point estimates from the dynamic IRT model by Martin and Quinn (2002) and the hidden Markov IRT model: Brennan, White, Frakfurter, and Blackmun. The chain was run for 20,000 draws after throwing out the first 10,000. Thick lines on the plots in the top two rows indicate posterior means and light lines are 95% Bayesian credible intervals. The plots in the bottom row show posterior probabilities of being in state 1 (light lines) and in state 2 (dark lines).

$$\theta_{\text{Stewart}} \sim N(1, 0.1),$$
$$\theta_{\text{Rehnquist}} \sim N(3, 0.1).$$

Also, these informative priors serve to limit the bounds of ideal point estimates; ideal point estimates near −2 and 3 are highly extreme values in this scale.

Figure 19.1 shows the results of the hidden Markov analysis of the 43 US Supreme Court justices. The fitted hidden Markov IRT model finds a break in ideal points for each justice. By checking the size of the break, we can tell whether a justice's preferences have actually changed.

Sixteen justices exhibit dramatic ideal points changes over their careers in the Court. Harlan, Black, Douglas, Marshall, Brennan, Warren, Frankfurter, Reed, Jackson, Blackmun, Rehnquist, Stevens, Souter, Thomas, O. J. Roberts, and Stone have dramatic shifts in their terms. However, significant preference changes are not found in Stewart, White, Whittacker, Burger, Kennedy, Scalia, Ginsburg, Breyer, Murphy, and O'Connor.

The results of the hidden Markov analysis identify substantively important issues about the timing and grouping of ideal point changes that may be indicative of broader social and political contextual factors. First, ideal points of the justices who served early in the sample period—Frankfurter, McReynolds, O. J. Roberts, and Stone—changed dramatically

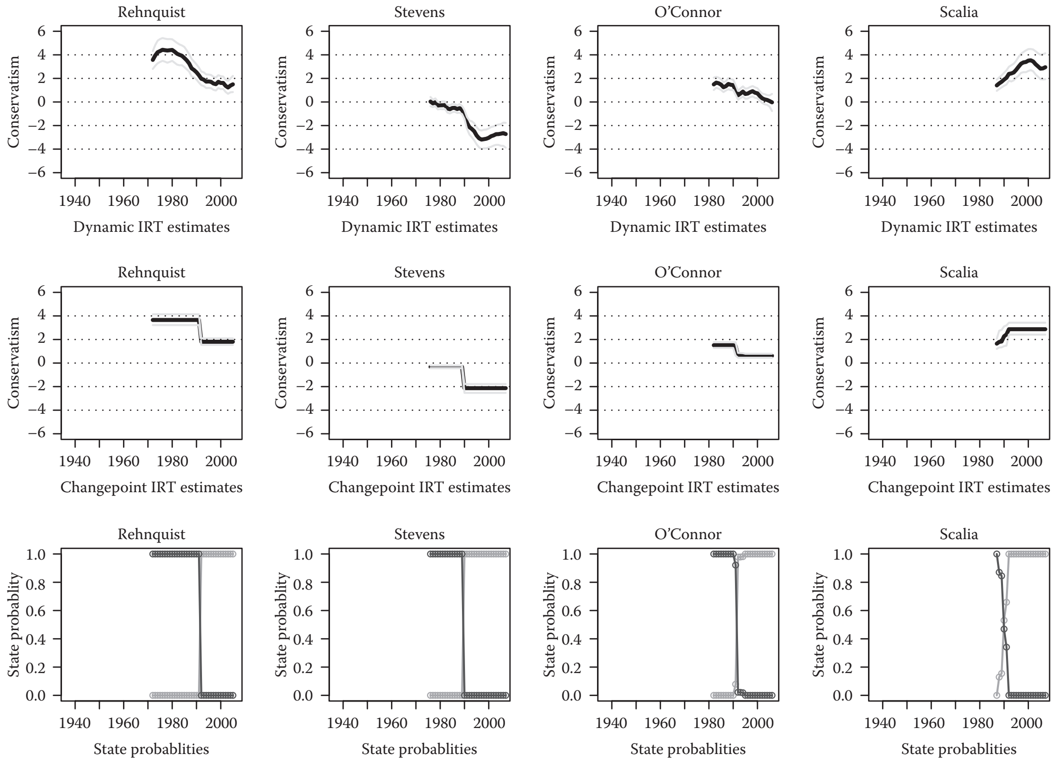

FIGURE 19.4
Comparison of ideal point estimates from the dynamic IRT model by Martin and Quinn (2002) and the hidden Markov IRT model: Rehnquist, Stevens, O'Connor, and Scalia. The chain was run for 20,000 draws after throwing out the first 10,000. Thick lines on the plots in the top two rows indicate posterior means and light lines are 95% Bayesian credible intervals. The plots in the bottom row show posterior probabilities of being in state 1 (light lines) and in state 2 (dark lines).

between the late 1930s and the early 1940s. Given the timing of the breaks, these preference changes seem likely to be related to what is known commonly as "the switch in time that saved nine," when Justice O. J. Roberts shifted his alignment to the liberal bloc of justices on a key 1937 case, a move that is often viewed as an means to protect the Court's independence from President Franklin Roosevelt's attempts to reorganize it through expansion (Epstein and Walker, 2007; Ho and Quinn, 2010).

Another interesting finding is timing of David Souter's preference shift. Souter, who was nominated by George H. W. Bush in 1990, has drawn the ire of conservatives for voting with liberal justices on many important cases including *Planned Parenthood* v. *Casey* and *Bush* v. *Gore*. When George H. W. Bush's son, George W. Bush, sought to fill two openings during his presidential term, conservatives fretted over whether his conservative picks would exhibit a similar leftward drift over their careers. The hidden Markov IRT model detects Souter's movement to the left in the early 1990s, very shortly after his confirmation. This movement was not so much a slow evolution, but a quick about-face followed by a long, consistent liberal preference.

Figures 19.2 through 19.4 compare the estimates from the hidden Markov IRT model with the estimates from Martin and Quinn (2002)'s dynamic IRT model. To save space, I select 12 justices with more than 20 years of service in the Court. In each figure, the top row replicates Martin and Quinn (2002)'s dynamic ideal point estimates, the middle row shows the hidden Markov ideal point estimates, and the bottom row shows the posterior regime probabilities from the hidden Markov IRT model.

The difference is clear. While Martin and Quinn's (2002) dynamic IRT model tracks trends in ideal points over time, the hidden Markov IRT model provides a sharp estimate of preference changes. For example, in the case of Douglas in Figure 19.2, it is hard to pinpoint the timing of Douglas's change from the Martin and Quinn (2002) estimate. By contrast, the hidden Markov IRT model clearly shows the timing of the shift. However, when justices' ideal points change slowly as in the case of Blackmun, the timing of the break uncovered by the hidden Markov IRT model is not as informative as the estimate from Martin and Quinn's (2002) dynamic IRT model.

19.5 Conclusions

In this chapter, I present a statistical model for dynamic ideal point estimation. The model combines the hidden Markov model with the standard two-parameter IRT model. The application of the the model to the US Supreme Court data demonstrates that the hidden Markov IRT model is an effective method to detect preference changes from longitudinal voting data.

Acknowledgments

The author would like to thank Andrew D. Martin and Kevin M. Quinn for providing the data and C++ codes for their dynamic IRT model; as well as John Balz, Andrew D. Martin, and Michael Peress for providing comments on the manuscript.

References

Albert, J. H. 1992. Bayesian estimation of normal ogive item response curves using Gibbs sampling. *Journal of Educational Statistics*, 17:251–269.

Albert, J. H. and Chib, S. 1993. Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association*, 88(422):669–679.

Bafumi, J., Gelman, A., Park, D. K., and Kaplan, N. 2005. Practical issues in implementing and understanding Bayesian ideal point estimation. *Political Analysis*, 13(2):171–187.

Bailey, M. and Chang, K. 2001. Comparing presidents, senators, and justices: Interinstitutional preference estimation. *Journal of Law, Economics, & Organization*, 17(2):477–506.

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. 2000. The statistical analysis of legislative behavior: A unified approach. Paper presented at the Annual Meeting of the Political Methodology Society.

Epstein, L. and Knight, J. 1998. *The Choices Justices Make*. CQ Press, Washington, D.C.

Epstein, L. and Walker, T. G. 2007. *Constitutional Law for a Changing America: Institutional Powers and Constraints*. CQ Press, Washington, DC.

Epstein, L., Hoekstra, V., Segal, J. A., and Spaeth, H. J. 1998. Do political preferences change? A longitudinal study of U.S. Supreme Court justices. *Journal of Politics*, 60(3):801–818.

Heckman, J. J. and Snyder, J. M. 1997. Linear probability models of the demand for attributes with an empirical application to estimating the preferences of legislators. *Rand Journal of Economics*, 28:142–189.

Ho, D. E. and Quinn, K. M. 2010. Did a switch in time save nine? *Journal of Legal Analysis*, 2(1):69–113.

Jackman, S. 2001. Multidimensional analysis of roll call data via Bayesian simulation: Identification, estimation, inference and model checking. *Political Analysis*, 9:227–241.

Johnson, V. E. and Albert, J. H. 1999. *Ordinal Data Modeling*. Springer, New York.

Londregan, J. 2000. Estimating legislators' preferred points. *Political Analysis*, 8(1):35–56.

Martin, A. D. and Quinn, K. M. 2002. Dynamic ideal point estimation via Markov chain Monte Carlo for the U.S. Supreme Court, 1953-1999. *Political Analysis*, 10(2):134–153.

Martin, A. D. and Quinn, K. M. 2007. Assessing preference change on the US Supreme Court. *Journal of Law, Economics, & Organization*, 23(2):365–385.

McCarty, N. M., Poole, K. T., and Rosenthal, H. 1997. *Income Redistribution and the Realignment of American Politics*. AEI Press, Washington, DC.

Poole, K. T. and Rosenthal, H. 1997. *Congress: A Political-Economic History of Roll-Call Voting*. Oxford University Press, Oxford.

Segal, J. A. and Spaeth, H. J. 1993. *The Supreme Court and the Attitudinal Model*. Cambridge University Press, Cambridge.