Estimating Ideological Locations in Australian Political Institutions

Simon Jackman*

September 25, 2001

*Department of Political Science, Stanford University, Stanford, California 94305-6044, USA. E-mail: jackman@stanford.edu

Roll Call Analysis

- Use the recorded votes of deliberative bodies to infer the "revealed preferences" of their members.
- "Deliberative bodies" includes courts, committees, legislatures.
- Goal: generate *measures* of legislators' preferences.
- Measures of are used in subsequent analyses of legislative politics: party cohesion, effects of party discipline, evolution of coalitions over time, dimensionality of the policy space.

Roll Call Analysis in non-Westminster settings

- Critical in the study of the U.S. Congress; literally hundreds of articles relying on various measures of legislative preferences

- Measures of legislative preferences used to:

  1. identify pivotal legislators: median legislators, filibuster pivots and veto pivots, the width of the "gridlock region"
  2. assess party cohesion
  3. effects of party switching
  4. committee assignments

Other Settings

- Historical analyses: party cleavages in the pre-Civil War U.S. Congress; structure of the Confederate Congress
- European Parliament: party loyalty versus voting as national blocs
- United Nations General Assembly
- Russian Parliament; nascent party system

"Real-world" uses of Preference Measures

- Interest-groups generate their own rankings of legislatures: e.g., Americans for Democratic Action (ADA), AFL-CIO, National Taxpayer's Union, Sierra Club, Chamber of Commerce, American Civil Liberties Union.

- Legislators themselves use these rankings to promote themselves as reliable conservatives or liberals; and to distinguish themselves from political opponents.

Is Roll Call Analysis Redundant in Westminster-Style Legislatures?

- Westminster legislatures characterized by

  1. executive drawn from the legislature and hence strong party discipline;
  2. single member districts; hence small number of parties

- Party discpline induces little or no variation in voting profiles for legislators of the same party

Is Roll Call Analysis Redundant in Westminster-Style Legislatures?

- If absolutely zero within-party variation, then each party can be treated as a *unitary actor*.

- If two perfectly disciplined parties, then only two unitary actors -- no unique scaling of the parties is possible (any two points will do, e.g., “left” and “right”).

Australian Legislatures

- Both Senate and House of Reps characterized by strong party discipline
- Roll call analysis for the House is uninteresting; save for issue of locating the (growing) number of independents?
- Senate a slightly more interesting case: method of election ensures minor party representation; Colston defection; occasional lapses of party loyalty

Questions

- Does statistical apparatus (motivated by the spatial model of voting) yield more than less formal approaches?

- Direct inspection of voting patterns

  1. Brown votes with the coalition: 7/52, 13%
  2. Democrats vote with the coalition: 15/55, 27%.
  3. Harradine (IND-TAS) votes with the coalition: 15/36, 44%
  4. Labor and the coalition vote together to defeat Green or Democrat proposals: 25/55, 45%.
  5. Harris (QLD-PHON) votes with the coalition: 16/24, 67%

- What is the dimensionality of the policy space?

Data: Australian Senate, 2001

- All recorded divisions in the Senate, for calendar year 2001 (up through Sept 20); gathered from *Journals of the Senate* and *Hansard*
- $n = 77$; Cherry (DEM, QLD) replaces Woodley, but have non-overlapping voting histories for both.
- $m = 55$ votes. Through September 21, the U.S. Senate has had 284 roll calls.
- High rates of missing data (see figure).
- 3,245 individual "Ayes" and "Noes" being modeled

Data

- Two rare lapses of party discipline, both by Democrats:
  1. May 23: request for government documents relating to HIH Insurance; passed 33-32, with all Democrats except Murray (WA) voting Aye
  2. June 28: Democrats split 3-5-1 on the third reading of the Interactive Gambling Bill 2001 (passed 34-28).

- No lapses of party discipline among ALP or Coalition.

“Measure with a Model”

- Use a Euclidean spatial voting model to analyze these data
- Contrast other approaches, such as factor analysis etc.
- Factor analysis not well suited for the analysis of binary data, and missing data.

The Euclidean Spatial Voting Model

- Legislators: $i = 1, \ldots, n$
- Roll Calls: $j = 1, \ldots, m$
- Data:

$$y_{ij} = \begin{cases} 1 & \text{legislator } i \text{ votes ``Aye'' in } j\text{-th division} \\ 0 & \text{legislator } i \text{ votes ``No'' in } j\text{-th division} \\ \text{NA} & \text{all forms of abstention} \end{cases}$$

- $\mathbf{Y} = \{y_{ij}\}$, a n by m matrix of individual voting decisions

Spatial Voting Model

- each legislator has an "ideal point" $\mathbf{x}_i$, a location in Euclidean space. In one-dimension the issue space is the left-right ideological continuum.

- each recorded vote is a choice between a proposal $\boldsymbol{\theta}_j$ and a reversion/status-quo point $\boldsymbol{\psi}_j$

- random utilities defined for each outcome, with quadratic loss:

$$u_i(\boldsymbol{\theta}_j) = -|\mathbf{x}_i - \boldsymbol{\theta}_j|^2 + \eta_{ij}$$
$$u_i(\boldsymbol{\psi}_j) = -|\mathbf{x}_i - \boldsymbol{\psi}_j|^2 + \nu_{ij}$$

Spatial Voting Model

y_{ij}^* denotes the latent utility difference between the proposal and status quo positions for the *i*th legislator,

$$y_{ij}^* = u_i(\boldsymbol{\theta}_j) - u_i(\boldsymbol{\psi}_j)$$

$$y_{ij}^* > 0 \iff y_{ij} = 1 \iff \text{“Yea”}$$

$$y_{ij}^* \leq 0 \iff y_{ij} = 0 \iff \text{“Nay”}$$

Statistical Model

Substituting for the utilities and re-arranging,

$$
\begin{aligned}
y_{ij}^{*} &= u_i(\boldsymbol{\theta}_j) - u_i(\boldsymbol{\psi}_j) \\
&= -|\mathbf{x}_i - \boldsymbol{\theta}_j|^2 + |\mathbf{x}_i - \boldsymbol{\psi}_j|^2 + \eta_{ij} - \nu_{ij} \\
&= 2\mathbf{x}_i'(\boldsymbol{\theta}_j - \psi_j) - |\boldsymbol{\theta}_j|^2 + |\boldsymbol{\psi}_j|^2 + \eta_{ij} - \nu_{ij} \\
\frac{y_{ij}^{*}}{\sigma_j} &= \mathbf{x}_i'\boldsymbol{\beta}_j - \alpha_j + \varepsilon_{ij}
\end{aligned}
$$

i.e., a **latent** linear regression model, where

$$
\begin{aligned}
\boldsymbol{\beta}_j &= 2(\boldsymbol{\theta}_j - \boldsymbol{\psi}_j)/\sigma_j \\
\alpha_j &= (\boldsymbol{\theta}_j^2 - \boldsymbol{\psi}_j^2)/\sigma_j \\
\varepsilon_{ij} &= (\eta_{ij} - \nu_{ij})/\sigma_j \\
\sigma_j^2 &= V(\eta_{ij}) - 2C(\eta_{ij}, \nu_{ij}) + V(\nu_{ij}) = 1
\end{aligned}
$$

A Probit Model

Assume $\varepsilon_{ij} \sim N(0, 1)$, $\forall\, i, j$. Then the probability of a "Yea" vote is

$$\begin{aligned} Pr(\text{“Yea”}_{ij}) &= Pr(y_{ij}^{*} > 0) \\ &= Pr\left(\mathbf{x}_i'\boldsymbol{\beta}_j - \alpha_j + \varepsilon_{ij} > 0\right) \\ &= \Phi\left(\mathbf{x}_i'\boldsymbol{\beta}_j - \alpha_j\right), \end{aligned}$$

where Φ is the standard normal CDF.

- This is a probit model, but with a significant complication: everything on the right-hand side of the model is **unobserved**.

- That is, we want estimates of **both** the bill parameters $(\boldsymbol{\beta}_j, \alpha_j)'$ and the unobserved "covariate" $\mathbf{x}_i$.

Meanwhile, in psychometrics…

In the educational testing literature, this model is known as a **two-parameter item-response model**.

$$Pr(\text{“Correct Answer”}_{ij}) = \Phi\left(x_i\beta_j - \alpha_j\right)$$

- The slope parameter β_j is an *item discrimination* parameter
- The intercept α_j is known as the *item difficulty* parameter
- x_i is the latent ability of the *i*-th test-taker

Estimation via Maximum Likelihood Is Usually Intractable

With n legislators, m roll calls and a d dimensional policy space, direct MLE is a $nd + (d + 1)m$ dimensional optimization problem

			Dimensions (d)		
	n	m	1	2	3
105th U.S. Senate	100	534	1,168	1,802	2,436
93rd U.S. House	442	917	2,276	3,635	4,994
U.S. House, 1789-1985	9,759	32,953	75,485	118,017	160,549
U.S. Senate, 1789-1985	1,714	37,281	76,276	115,271	154,266
Australian Senate, 2001	77	55	187	319	451

Estimation via Bayesian Simulation

- Moreover, the model is unidentified due to *scale invariance* -- require constraints for unique set of estimates

- Switch to Bayesian methods: prior distributions for all parameters parameters; in particular, $x_i \sim N(0, 1) \, \forall \, i$ provides a reference scale.

- Sample repeatedly from the *posterior distribution* for the model parameters, by sampling from lower-dimensional conditional distributions

One-Dimensional Model

- Fits extremely well, especially for the major parties (see figure)
- With a classification threshold of 0.5, 92.2% of 3,245 votes correctly predicted.
- See lack-of-fit figure
- Notable lack of fit for Democrats, Harradine and Harris (QLD, PHON).

Rank Ordering

Because we estimate the joint density of the ideal points for all legislators, we can perform inference: in particular, we can test conjectures about the recovered rank ordering

- Faulkner $<$ Cooney: $p = .78$
- Cooney $<$ Brown: $p > .99$
- Cooney $<$ Harradine: $p > .99$
- Brown $<$ Harradine: $p = .81$
- Harradine $<$ Stott Despoja: $p = .97$
- Stott Despoja $<$ Murray: $p = .68$

Rank Ordering

- Woodley $<$ Cherry: $p = .55$
- Harris $<$ Ian Macdonald: $p > .97$
- Ian Macdonald $<$ Tchen: $p = .73$
- The Senate Median is a Democrat: $p > .99$.

Two-Dimensional Model

- Percent correctly classified goes up to 99.4%
- Harris 84.7%; Harradine 89.1%; Brown 97.5%
- Poorest classification by division, 92.7%, Interactive Gambling Bill (3rd reading, Dems split).
- Breakdown of divisions:

18 (33%)	purely "left-right", with Democrats pivotal
25 (45%)	purely "vertical", Democrats vs major
12 (22%)	mix of left-right, up-down.

- Visualization of feasible policy region (see figure)