A Latent Class Framework for Roll Call Analysis

Juraj Medzihorsky

V-Dem Institute, University of Gothenburg

Abstract

This paper introduces to roll call analysis a latent class framework that abandons the standard assumption that the model holds in the whole population. The framework splits the population into a class that belongs to the model and a minimized residual class, and draws inferences from the size and composition of the latter. The framework generalizes two widely used measures, the Rice cohesion index and the party unity score, and generates new methods that identify votes inconsistent with group cohesion defined in terms of ideal points, or with non-partisan voting. In ideal point analysis, the framework also aids validation and substantive interpretation with a differential item functioning diagnostic. The methods are demonstrated on congressional roll calls on the Civil Rights Act of 1964.

The Mixture Index of Fit

The π^* mixture index [1] measures the misfit of a model by the smallest fraction of the population that it cannot describe,

$$\pi^*(X, \mathcal{M}) = \inf\{\pi : X = (1-\pi)\,M + \pi\,R,\ \pi \in [0,1],\ M \in \mathcal{M},\ R \text{ unspecified}\}, \tag{1}$$

where π^* is the size of the undescribed fraction, X is the observed distribution, M an element from the model $\mathcal{M}$, and R an unspecified residual distribution. The index rests on always true assumptions, is easy and intuitive to interpret, and provides a new kind of substantively informative residual analysis. The mixture index framework extends the toolbox of legislative scholars in two ways. First, it allows to define new measures that are easy to interpret and based on always true assumptions. These measures rest on comparing the data with a substantively interesting baseline model, and capturing their distance as the smallest fraction of the observations that cannot be described by the baseline model. This paper presents such measures of group cohesion and partisan voting. Second, it provides a differential item functioning diagnostic that can aid the validation and substantive interpretation of ideal point estimates.

Ideal Point Group Cohesion

From the ideal point perspective, a group can be cohesive even if its legislators do not always vote the same, provided they have similar ideal points. This notion of cohesion can be measured with the mixture index in two different ways. Under the first, the legislators are sorted into those who share an ideal point and those who do not. The second represents a more relaxed notion of cohesion, dividing the legislators into those with ideal points from the same restrictive distribution and those from an unrestricted one. The shared ideal point or the parameters of the restrictive distribution are estimated in the process. Figure 1 shows an example of both, using binomial and beta-binomial voting, respectively [2, 3].

		Votes with S. Thurmond: 0	1	2	3	4	5	6
	Observed	33	8		2	2	1	21
Binomial Voting	Model	33	1					
	Residual		7		2	2	1	21
Beta-Binomial Voting	Model	33						21
	Residual		8		2	2	1	

Figure 1: The group-fixed and Beta distributed ideal points models applied to the data on voting among the Democratic senators with S. Thurmond in six selected roll calls on the Civil Rights Act of 1964.

Group Voting

The mixture index provides an intuitive measure of group voting that rests on splitting the votes into those independent of the group and those associated with it, maximizing the size of the former class. This measure applies to any number of options, groups, as well as multiple potentially related group memberships. For any number of options and parties non-partisan voting corresponds to the log-linear model of independence

$$\ln m_{op} = \lambda^0 + \lambda^O_o + \lambda^P_p, \tag{2}$$

where m_{op} are the non-partisan votes for option o in party p, λ^0 the main term, and λ^O and λ^P the option and party terms, respectively. The smallest possible fraction of non-partisan votes is computed using this model as $\mathcal{M}$ in Equation 1. Figure 2 illustrates this approach with two decompositions of the final House passage of the CRA. The first splits the votes into those independent of the party and those associated with it and the second does the same for the region. Figure 4 contrasts the π^* measure of partisan voting with the difference between party unity scores on 120 CRA-related Senate roll calls, highlighting five patterns of party unity and partisan voting.

Non-Partisan Voting

	Model: Nay	Model: Aye	Residual: Nay	Residual: Aye
Democrat	39	153	52	
Republican	35	136		

Non-Regional Voting

	Model: Nay	Model: Aye	Residual: Nay	Residual: Aye
North	32	281		
South	1	8	93	

Figure 2: Two minimum mixture decompositions of the final House passage of the CRA.

	Non-Regional: North Nay	North Aye	South Nay	South Aye	Regional: North Nay	North Aye	South Nay	South Aye
Democrat			83					8
Republican	16		11		16			

Figure 3: Residuals under two models of non-partisan voting in the final House passage of the CRA.

The approach extends easily to multiple group memberships. For the case of party and region

$$\ln m_{opr} = \lambda^0 + \lambda^O_o + \lambda^P_p + \lambda^R_r + \lambda^{PR}_{pr}, \tag{3}$$

where λ^P and λ^R are the party and region terms, respectively, and λ^{PR} the party-region interaction accounting for different party shares in the regions. Under this model voting is neither partisan nor regional. Partisan voting can be introduced with the option-party interaction λ^{OP} and regional voting with the option-region interaction λ^{OR}. Figure 3 shows the application of two such baseline models–non-partisan non-regional voting, and non-partisan regional voting–to the final House passage of the CRA.

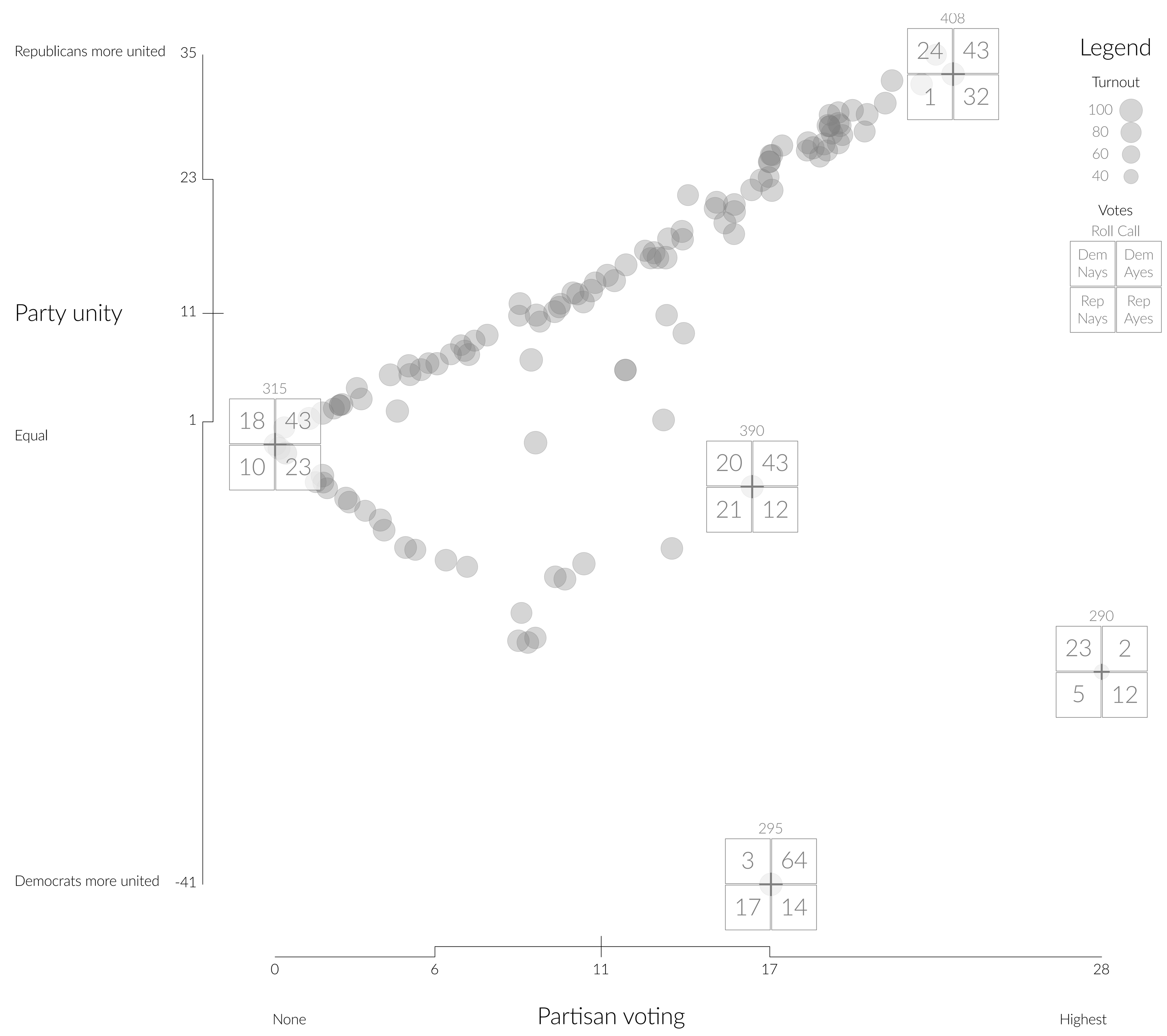

Figure 4: Partisan voting and difference in party unity for 120 Senate roll calls related to the Civil Rights Act of 1964. The axes show the minima and maxima, and 25 per cent, 50 per cent, and 75 per cent quantiles.

Differential Item Functioning Diagnostics for Ideal Point Scaling

In a differentially functioning roll call legislators with similar ideal points will vote differently conditional on some attribute of theirs such as party or gender. Thus, DIF detection can aid ideal point scaling. Roll call data is typically small and covers the whole population, limiting the usefulness of NHST-based diagnostics. The minimum mixture DIF diagnostic [4] is free of these drawbacks and identifies DIF-affected responses. It rests on two nested log-linear models. Under the first, no DIF, model the responses are associated with the ideal points, but not with the attribute. The second model allows a uniform association of the attribute and the response over the latent dimension. Figure 5 illustrates the first model.

Quantile on the First Dimension of DW-NOMINATE

		1st Nay	1st Aye	2nd Nay	2nd Aye	3rd Nay	3rd Aye	4th Nay	4th Aye	5th Nay	5th Aye
No DIF	North		79	2	62	6		5	57	19	55
	South		3		2	46	3				
Residual	North						28				
	South	1		17				21		9	

Figure 5: DIF diagnostics for the first DW-NOMINATE dimension [5] in the final House passage of the CRA. Decomposition into a DIF-free and residual components.

Discussion

- Interval estimates available via jackknife [6] and bias-corrected bootstrap [7].
- The mixture index has been generalized to missing data [8].
- Smoothing sparse roll call data to facilitate the application of the mixture index benefits from sensitivity analysis.
- Latent class IRT models [9] facilitate the application of the mixture index in roll call scaling, but this is currently computationally too demanding for general use.
- Computation with the `pistar` package [10], with EM nested in line-splitting or Brent root search.
- Ongoing work on frequentist and Bayesian computation of the mixture index with Hamiltonian Monte Carlo.

References

[1] Tamás Rudas, Clifford C. Clogg, and Bruce G. Lindsay. "A new index of fit based on mixture methods for the analysis of contingency tables". In: *Journal of the Royal Statistical Society. Series B (Methodological)* 56.4 (1994), pp. 623–639.

[2] Bradley Palmquist. "Heterogeneity and dispersion in the beta-binomial model". In: *the Annual Meeting of the American Political Science Association. Washington DC*. 1997.

[3] Scott W. Desposato. "Correcting for small group inflation of roll-call cohesion scores". In: *British Journal of Political Science* 35.4 (2005), pp. 731–744.

[4] Tamás Rudas and Rebecca Zwick. "Estimating the importance of differential item functioning". In: *Journal of Educational and Behavioral Statistics* 22.1 (1997), pp. 31–45.

[5] Nolan M McCarty, Keith T Poole, and Howard Rosenthal. *Income redistribution and the realignment of American politics*. AEI Press, 1997.

[6] C. Mitchell Dayton. "Applications and computational strategies for the two-point mixture index of fit". In: *British Journal of Mathematical and Statistical Psychology* 56.1 (2003), pp. 1–13.

[7] Jenő Reiczigel et al. "Bias-corrected estimation of the Rudas–Clogg–Lindsay mixture index of fit". In: *British Journal of Mathematical and Statistical Psychology* (2017).

[8] Tamás Rudas. "Mixture models of missing data". In: *Quality & Quantity* 39.1 (2005), pp. 19–36.

[9] Bruce G. Lindsay, Clifford C. Clogg, and John Grego. "Semiparametric estimation in the Rasch model and related exponential response models, including a simple latent class model for item analysis". In: *Journal of the American Statistical Association* 86.413 (1991), pp. 96–107.

[10] Juraj Medzihorsky. *pistar: Rudas, Clogg and Lindsay Mixture Index of Fit*. `https://github.com/jmedzihorsky/pistar`. R package version 0.5.2.5. 2015.

Acknowledgements

I am grateful to Tamás Rudas, Gábor Tóka, Jason Wittenberg, Jack Blumenau, Constantin Manuel Bosancianu, Federico Vegetti, Mihail Chiru, and Zsolt Enyedi for comments on earlier versions of this project. This research project was supported by European Research Council, Grant 724191, PI: Staffan I. Lindberg, V-Dem Institute, University of Gothenburg, Sweden.