

The model used for this is a Bayesian Dynamic-Ordinal Item Response Theory Model. This model is the same as that developed by ?. Rather than examining countries' respect for human rights, this model examines the level of *de jure* judicial independence that a state's judicial institutions represents.¹

I assume that the observed indicators for each state-year are functions of a single dimensional latent variable (the Judicial Independence Continuum) that represents the level of *de jure* judicial independence. For each state-year observation, let i index the state and t index the year. or each model, there are J indicators $J = 1, \dots, J$ each of which is ordinal.² I estimate each θ_{it} , which is the latent level of *de jure* judicial independence of each state i in year t (?, 7).

Let $i = 1, \dots, N$ index cross-sectional units and $t = 1, \dots, T$ index time periods. In each time, period, I observe values y_{ij} for each of $j = 1, \dots, J$ indicators for each unit. Each indicator is assumed to be ordinal and can take on K_j values.³ The responses to each of the items depend on a single latent variable θ_{it} , which may vary across units and over time. I assume that all indicators are independently drawn from a logistic distribution (?, 7).

For each indicator, there is an item discrimination parameter β_j and a set of $K_j - 1$ difficult cut-points.⁴ (?, 7). A benefit of using institutional arrangements, which are observable with readily available data is that any error that is introduced into the indicators is the result of coding errors rather than perceptual errors found in survey responses.

The probability distribution for a given response to item j is observed as:

$$P[y_{ij} = k] = F(\alpha_{jk} - \theta_{it}\beta_j) - F(\alpha_{j(k-1)} - \theta_{it}\beta_j) \quad (1)$$

$F(\cdot)$ is the logistic cumulative distribution function. I assume the local independence of responses across units, the likelihood function for β , α , and θ given the data is⁵:

$$\mathcal{L}(\beta, \alpha, \theta|y) = \prod_{i=1}^N \prod_{t=1}^T \prod_{j=1}^J [F(\alpha_{jy_{it}} - \theta_{it}\beta_j) - F(\alpha_{jy_{it}-1} - \theta_{it}\beta_j)] \quad (2)$$

¹There is also one major difference in that ? measure a *de facto* level of respect for human rights, in that they look at actual numbers of torture, extra-judicial killings, and disappearances in a given year, I am looking at institutional arrangements, which typically has much less within-state variance from year to year.

²For definitions of each level of indicators, see Table ??

³For term length indicators, these have been collapsed into five ordered categories. This is due to the relative infinite value of a life or quasi-life term. A previous specification of this model was considered using the average length of term in office for each justice that had served on the court under the life or quasi-life term. This specification was rejected as introducing a questionable degree of *de facto* independence into what is a purely *de jure* measurement model.

⁴For more detail on these cut-points see (??).

⁵For a more detailed explanation of the assumptions made in this likelihood, see (?, 8).

If θ , (*de jure* judicial independence), was able to be observed, the likelihood function shown in Equation 2 would be equivalent to an independent ordinal logistic regression model (2, 8). There are three basic independence assumptions considered for IRT models. The assumptions are: (1) local independence of different indicators within the same state – year, (2) local independence of indicators across states within years and (3) local independence of indicators across years within states (2, 8). The implication of this is that each y_{itj} are independently drawn conditional on θ . The only relationship between two item responses is that they are both measure the same latent variable, θ . 2, 8 discuss the assumptions made in this model in greater detail. I relax Assumption 1 slightly in that there is some correlation between selection and retention methods, as well as between term lengths. These correlations are shown in Appendix ???. For example, there are no states which use a commission system for initial selection and then require judges to run in competitive elections.⁶ However, their Assumption 3, is a matter of concern in this model for the same reason as (?). Assumption 3 relates to the local independence of indicators across years within states (2, 8). Rather than attempting to measure a regime’s respect for human rights as in 2, I am measuring an institutional arrangement by each state. In each state there is a yearly potential to change selection methods, keeping intact a relaxed version of this assumption. This relaxation of Assumption 3 comes from the informative priors used by 2, 8.

⁶This assumption does not extend to filling judicial vacancies. For example, New Mexico uses a merit selection system to fill judicial vacancies and then requires the incumbent to run in a competitive partisan election at the next general election (?). I consider this to be an artifact of the *de facto* nature rather a factor of *de jure* judicial institutions.