Estimating Ideology of Brazilian Legislative Parties, 1990–2005: A Research Communication Web Appendix

Timothy Power
Latin American Centre
Oxford University
timothy.power@brazil.ox.ac.uk

Cesar Zucco Jr.
IUPERJ - Instituto Universitário de Pesquisas do Rio de Janeiro
czucco@iuperj.br

January 28, 2009

1 Standard Errors of Party Estimates

For each survey, standard errors for the party estimates π_k are obtained directly from the maximum likelihood procedure. There is no intrinsic metric for these estimates, so as we computed a rescaling factor that converts the *individual* legislators's rescaled placements back into the 1-10 scale adopted in the surveys. We then applied the same scale to the party estimates. This is a linear transformation of a random variable, and hence requires the correspondent transformation its standard errors. Assuming that the transformation of the estimates into the 1-10 scale is is obtained by adding (or subtracting) a and multiplying the result by b, the standard error of these estimates in the new scale is $b\times$ the original standard error.

A related, but significantly more important issue occurs when computing the standard errors of the party positions that incorporate year effects. Call π_{jt}, the estimates of the position of party j in a given year t, which has a standard error σ_π. The double rescaled positions are $\pi^*_{jt} = \frac{\pi_{jt}-\gamma_t}{\delta_t}$, but now the rescaling factors have are also estimates, and thus have standard errors σ_δ and σ_γ.

One way to compute the standard error of π^*_{jt} is to apply the delta method, which in this case consists of pre and post multiplying a vector of the derivatives of π^*_{jt} with respect to each of the three random variables by the variance-covariance matrix of these estimates.

The problem, is that π_{jt} is not estimated in the same procedure as δ and γ, and it is not obvious what — if any — would be the covariance structure between these terms.

$$VAR(\pi^*_{jt}) = \begin{pmatrix} \frac{1}{\delta_t} & -\frac{\gamma-\pi_{jt}}{\delta_t^2} & -\frac{1}{\delta} \end{pmatrix} \times \begin{pmatrix} \sigma_\pi & ? & ? \\ ? & \sigma_\delta & \sigma_{\delta\gamma} \\ ? & \sigma_{\gamma\delta} & \sigma_\gamma \end{pmatrix} \times \begin{pmatrix} \frac{1}{\delta_t} \\ -\frac{\gamma-\pi_{jt}}{\delta_t^2} \\ -\frac{1}{\delta} \end{pmatrix} \tag{1}$$

As a practical solution, we assumed the unknown terms in Expression 1 to be zero. While not the optimal solution, this is a considerable improvement over treating δ and γ as constants. The resulting standard errors are somewhat larger for the larger parties (by an average factor of 1.4) and on average three times larger for the smaller parties.

In the paper, we provide the estimates incorporating both sources of uncertainty. As a consequence, if readers are interested in using only the estimates from a single survey, the standard errors will be considerably smaller for all years, except for 1990 which was used as the base year in the cross-year estimation. In the party data files, available online, we provide both sets of standard errors.

2 Treatment of Missing Data

The answers to the ideological placement questions in the surveys exhibit two different "types" of missing responses: either the actual legislator self placement or one or more placements of parties were missing. Though none was too pervasive, we opted to deal with both issues instead of simply dropping the observations because of incomplete data. In the few cases of the first type, we assumed legislators would have placed themselves where they placed their parties. Only if that information was missing too — because the legislator either did not state or place his/her party or — was the legislator dropped from the sample. *Missingness* in the party placements, on the other hand, was dealt with by use of a simple imputation algorithm. We first estimated the α_i, β_i, and π_j for the set of legislators with complete responses and computed average predicted values to use as seeds. We then re-estimated the parameters and replaced the seeds with predicted values until the estimates converged. Estimates always converged fairly quickly, and the estimates for the other legislators (those without missing data) were unchanged. Since they are in effect averages, the imputed estimates do not affect the estimation of the party position nor the individual rescaling parameters, while allowing us to use all the information effectively available.

3 The use of priors and truncation

In the results produced by the basic model, a few legislators' rescaled positions take on extreme values (See Figure 1(a) for an example from 1993), generally caused by unusual patterns of party placement. Such extreme positions are unrealistic, so I dealt with them by the imposition of "priors" on α_i's and β_i.

This was done by establishing a *prior distribution* from which the parameters are drawn. I adopted the common prior that all legislators do not distort the scale, which amounts to assuming all α_i's were drawn from a prior distribution with mean zero ($\alpha_\mathbf{i} \sim N(\mu_\alpha = 0, \sigma_\alpha = 1)$) and all β_i's are drawn from a prior distribution with mean equal to 1 ($\beta_\mathbf{i} \sim N(\mu_\beta = 1, \sigma_\beta = 1)$).

Priors have a simple interpretation in the maximum likelihood framework. If we interpret the probability of observing any given data as the posterior conditional distribution, we know that this is proportional to the product of likelihood and the prior, such that:

$$Pr(\alpha, \beta | P_{ij}) \approx Pr(P_{ij} | \alpha, \beta) \times Pr(\alpha, \beta) \tag{2}$$

Priors work as a *penalty* on the log likelihood function that prevents the estimates from deviating too much from the prior mean. As a consequence, the posterior mean will always lie between the prior and the data mean, with each distribution weighted by its precision. Hence, with the inclusion of these priors, the new maximum likelihood estimator for α and β becomes simply:

$$\begin{aligned}\mathcal{L}^* = \sum^{i}\sum^{j} - \log(\sigma) - \frac{1}{2\sigma^2}\left(P_{ij} - \alpha_i - \beta_j \pi_j\right)^2 \\ - \sum^{i} \log(\sigma_\alpha) - \frac{1}{2\sigma_\alpha^2}(\alpha - \mu_\alpha)^2 \\ - \sum^{i} \log(\sigma_\beta) - \frac{1}{2\sigma_\beta^2}(\beta - \mu_\beta)^2\end{aligned} \tag{3}$$

Recall that in the basic model, the likelihood function was simply

$$\mathcal{L} = \sum^{i}\sum^{j} - \log(\sigma) - \frac{1}{2\sigma^2}\left(P_{ij} - \alpha_i - \beta_j \pi_j\right)^2 \tag{4}$$

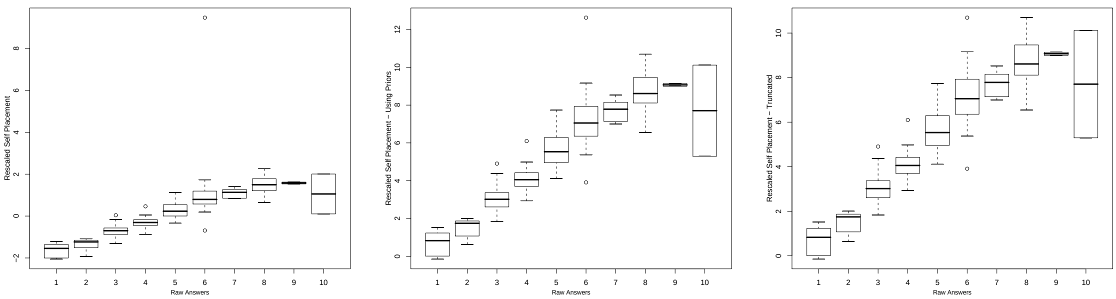
Hence, a quick examination of Eqs. 4 and 3 shows that in practical terms priors represent simply an extra set of parameters that are added to the likelihood function.

Only two outlying cases remained after the estimation with the use of priors both of which are worth considering in a little more detail.

- In the 1993 survey (Figure 1(b)), a legislator from the PPR/TO placed all parties very close together, and himself together with the rightmost party. Consequently, his rescaled position, even after imposing priors, was considerably higher than the next legislator.

- In the 1997 survey, a legislator from the PMDB/SC placed all parties in a small section of the scale in a rather "unorthodox" ordering. Since he placed himself on the edge of the section he used, as far as possible from the PCdoB (Communist Party of Brazil) his rescaled placement shot away to the right.

In both these cases the procedure captured the side of the scale the legislators should be in, but their answers forced them too far away, even after using the priors. Thus, as a last step in the rescaling procedure we truncated the scale at the last non-outlying legislator, and converted it to the original metric for comparison purposes. The results after each of these steps (simple estimation, estimation with priors, and with truncation of outliers) are shown for the 1993 survey in Figure 1.



(a) Simple Rescaling (b) Rescaling with Priors (c) Priors and Truncation

Figure 1: The Three Steps in the Rescaling Procedure — 1993 Survey

References