Formal Comparisons of Legislative Institutions: Comparing Ideal Points from the Chamber of Deputies and the Federal Senate in Brazil

Robert Myles McDonnell*
Fundação Getúlio Vargas

Abstract

As the popularity of formal analyses of legislative activity in Latin America grows, so does the importance of understanding the limits of the estimates produced by such analyses and the methodological adaptations necessary when using these measures to make formal comparisons. This research note details the considerations involved and demonstrates their significance with an empirical example using the Brazilian Chamber of Deputies and the Federal Senate. This empirical analysis leads to conclusions that are the *opposite* of those in the literature, suggesting that such formal comparisons across institutions need to be made with care.

Keywords: ideal points; legislative voting; Bayesian item-response; federal senate; chamber of deputies.

1. Introduction

When making formal comparisons of legislative arenas, there are methodological considerations that researchers must keep in mind. In this research note, I outline these concerns and demonstrate with an example how findings can change once these considerations are taken into account. By ‘formal comparison’, I mean quantitative estimates of legislative voting behaviour, such as ideal points. As the “primary use of roll call data [...] is the estimation of ideal points” (Clinton et al. 2004, 1), I focus on these particular estimates

*The author would like to thank Danilo Freire and Guilherme Duarte for suggestions for changes to this paper.

throughout this research note. The empirical example I use is a comparison of ideal point estimates from the Brazilian Chamber of Deputies and Federal Senate.

2. Comparisons across institutions

Calling attention to the difficulties in making formal comparisons across legislative arenas or similar institutions is not new. [Bailey \(2007\)](#) and [Treier \(2011\)](#) both highlight the misleading inferences that may result from comparisons of this sort when scholars do not include the methodological adaptations necessary. In fact, these adaptations are conceptually very simple. If we consider the basic Bayesian Item-Response Theory (IRT) model, employed in this context by [Clinton et al. \(2004\)](#) and others, we can see clearly the basic requirements for making valid comparisons of different institutions. This model may be written as:

$$P(y_{ij} = 1) = \Phi(\beta_j\theta_i - \alpha_j)$$

where y_{ij} is a matrix of recorded nominal votes, with i indexing legislators and j votes, in which $y_{ij} = 1$ indicates that legislator i voted ‘yes’ on proposal j , and $y_{ij} = 0$ otherwise. β is the ‘discrimination’ parameter of the model and α the ‘difficulty’ parameter; Φ indicates a probit or logistic link function, and θ_i is the ideal point of legislator i , which may be understood as his/her preferred location in the policy space, a space of usually of one or two dimensions ([Jackman 2001](#)). In regression terms, α is the intercept and β the slope; this is a latent regression model, as we only observe y_{ij} .

This model originally comes from the literature on educational tests, the context of which makes it very clear what is necessary in order to produce comparable ideal points. The model is designed to gauge the worth of an educational test, for example, on mathematical ability. The questions may range in terms of difficulty, and if individuals of varying aptitude in the latent trait (mathematical ability) answer the same questions in the same manner, these questions are unable to discriminate between the test-takers. That is, test-takers who have different levels of ability will have the same probability of answering the question correctly for questions with low discrimination parameter values. Questions that can separate individuals of differing capacity do discriminate between test-takers, and result in the model predicting that individuals of greater ability have a higher probability of answering correctly. Although ideal points are not usually the main items of interest in this literature, with this model we can recover estimates of where the test-takers are positioned on the latent trait scale. For tests that discriminate well, we would expect to find those of limited mathematical skill at one end of the scale, while those with higher levels should be found at the other.

In the present context, the scale is constructed from the voting behavior of the legislators as compared to one another; two legislators who never vote the same way will be on the two extremes of the scale, for example, whereas those who vote similarly will be placed along the same region. The positions of the legislators on the scale are their ideal points. High absolute values of the discrimination parameter (above 2 or below -2) indicate that the proposal separated the legislators; in other words, all the legislators do not vote in the same

way for proposals with these high values. Conversely, values of this parameter that are close to zero indicate that all or most of the legislators voted the same way on such proposals.

It is obvious from this short description that test-takers must answer the same questions in order for their positions on the scale to be comparable. In the political context, we may not have the same control over what ‘questions’ (votes) legislators respond to, but this requirement remains the same. Hence, to compare two different institutions, both sets of legislators must respond to the same bills. Depending on the specific context, this may or may not even be possible. It also may not be justifiable: legislators in different houses may vote on very different versions of the same bill, for example, and bills may be amended many times before arriving in one house from the other. With regard to time, analyses of ideal points usually take a single legislature as a unit of analysis. Extending the period of time may necessitate the need for a dynamic model, such as that of [Martin and Quinn \(2002\)](#).

To expand upon this theme, I use the example of a comparison of ideal points between the Chamber of Deputies and the Federal Senate in Brazil, with a dataset from CEBRAP.¹ These two houses have been formally compared by [de Freitas et al. \(2012\)](#), [Bernabel \(2015\)](#) and [Desposato \(2006\)](#). Desposato utilizes the fact that deputies and senators vote sequentially in the National Congress to place both sets of legislators in a common space. This is the most valid method for doing so, and cleverly exploits a native feature of the Brazilian political system. However, as Bailey notes, votes must be coded as being the same item, as if they were the exact same question in an educational-testing context ([Bailey 2007](#), 440); there is no evidence available that [Desposato \(2006\)](#) undertook this necessary step. [de Freitas et al. \(2012\)](#) estimate ideal point means for both houses separately and then compare the means, which is not a valid method of formal comparison; Bernabel estimates ideal points for both houses separately, which is also invalid as a method of comparison, and although the study is offered as a descriptive piece ([Bernabel 2015](#), 106), this comparison is potentially quite misleading. For example, [Desposato \(2006\)](#) finds less party cohesion in the Senate when compared to the Chamber, whereas [Bernabel \(2015\)](#) finds the opposite – how are we to know which finding is correct?

To productively compare these houses, we must place both on a common scale, as noted by [Bernabel \(2015\)](#). In order to do so, we must know which proposals were voted on in both houses, and on which versions of the proposals the houses voted, and which are common to both houses. Exactly how to do so will differ on the specific context. In the current example, we can exploit the common votes in the National Congress and use these votes as a means to ‘bridge’ between the two houses. These votes are coded as being the same item in the database and therefore the two sets of legislators are linked by their voting on these common votes. The analysis may then proceed using all the remaining votes from both houses in one common dataset. This method is the same as that proposed by [Desposato \(2006\)](#), but leads to opposite findings. As mentioned previously, there is no evidence that Desposato treated the common votes in the National Congress as the exact same item, and I can only surmise that the difference in the findings presented here is down to this important detail.

¹CEBRAP For the data and code to reproduce the analysis done in this paper, see [here](#).

3. Data

The specific set of votes that were used to bridge the houses is that of the *Medida Provisórias* (MPVs, Provisional Decrees) issued by the Executive voted on in the National Congress, and is shown in Table 2 in the appendix. Many of these votes happened sequentially in the National Congress, and in a few cases the vote in the Senate happened at a later date (see Table 2). However, given that the voting is sequential in the National Congress (Desposato 2006), it is therefore not perfectly simultaneous, and this difference in time should not make any obvious difference. The Senators are aware of the result of the Deputies' votes in either case, and while it may be possible that the external political context shifted so rapidly as to change the probability of a 'yes' vote in the second house, this is unlikely for the small number of MPVs involved in the years analyzed (1989-2010). All other non-unanimous votes in both houses were then used in the normal manner, from one large conjoined dataset, once the bridging votes were coded as such.

There are of course other ways to bridge legislatures. One method is to use the legislators themselves, as in Shor et al. (2010). In the Brazilian case, voting behavior could be analyzed over several levels of government – from municipalities to federal legislatures, and as such is an interesting avenue for future research, but not the focus of the present research note. Another method is to use the content of the votes themselves and thus code votes on the same proposal that occur in different time periods as a means to bridge over time, as in Bailey et al. (2015); future studies on the dynamics of voting in both the Chamber and Senate could utilize this method.

Once we place the deputies and senators in a common space, we can meaningfully analyze hypotheses of differences between the institutions, operationalized by examining ideal point estimates from this common scale. Due to the large number of deputies, party means of ideal points instead of individual ideal points are considered below for both houses. Claims in the literature lead to several hypotheses, the first being, as argued by Desposato, *H1: there is no significant difference in voting behavior across the two houses*. This can be operationalized by examining the difference in means of the ideal points in both houses and the magnitude of the difference through a Bayesian version of the t-test for two groups (Kruschke 2013).

Some have posited differences in parties across the two legislatures, in particular that the PMDB, the *Partido do Movimento Democrático Brasileiro*, has differing factions depending on the house (Zucco Jr. and Lauderdale 2011). Our second hypothesis, operationalized in the same way as H1, is thus *H2: there is a significant difference in party voting behavior across houses, in particular the PMDB*.

Desposato (2006) and Bernabel (2015) also disagreed on party cohesion within these legislatures, with the former finding more cohesion in the Chamber and the latter finding the opposite. This suggests a third hypothesis, *H3: parties in the Chamber are more cohesive than those of the Senate*. In order to examine this hypothesis, I will employ the same method as before, but with the standard deviations of party ideal points in the houses rather than party mean ideal points.

It may also be the case that the Senate votes differently owing to its purpose (Neiva

and Soares 2013), its specific powers (Neiva 2008), the senior and perhaps conservative profile of the senators themselves (many are former presidential candidates, governors, party leaders and ex-ministers), or the high level of un-elected *suplentes* (stand-ins) in the Senate (Neiva and Izumi 2012); the hypotheses above will allow to establish if there is a difference in voting behavior between the Senate and the Chamber and its magnitude. These questions of why there is a difference are an important area for further research. In order to test these hypotheses, I utilize methods detailed in Kruschke (2013), implemented in **R** (R Core Team 2015) with the BEST package Kruschke and Meredith (2015). In the following discussion, I use the terms ‘left’ and ‘right’ as commonly understood, although the policy space could of course be another scale, such as government–opposition (Izumi 2011).

4. Results

Figure 1 shows the difference in means of ideal points for the two houses across all six legislatures studied. In each case, the Chamber is ‘mean 1’ (μ_1). The histograms show the posterior probability distribution of the difference of the mean ideal point of the Chamber minus that of the Senate, with the 95% Highest Density Interval (HDI) displayed at the base of the histogram. Were the differences zero in any legislature, we would see that some part of any one of the histograms touches zero. As can be clearly seen, not only is this not the case, but all the distributions are quite some distance from zero. Hence Hypothesis 1 finds no support – voting behavior in the two houses is quite clearly different, with only Lula’s second term in the 53rd legislature coming anywhere close to zero.

To assess the magnitude of this difference, we can examine the effect size of the difference in these mean ideal points, which is the difference in the means of the two groups, taking into consideration the variation within the groups.² Figure 2 shows effect sizes for the six terms considered. Legislatures 48, 49 and 51 – the presidencies of Sarney, Collor, Franco and Cardoso – have effect sizes of close to 0.7, meaning that there is roughly a 76% probability that a Deputy from these periods was to the right of a Senator from the same period.³ Legislature 52, Lula’s first term as president, displays an effect size of 0.955, meaning a probability of 82% that a Deputy in this term was more to the right than a Senator. The smallest effect size is to be found in Lula’s second term (a roughly 62% probability), whereas the largest, an enormous -3.06, is found in Cardoso’s first term, in which it is virtually guaranteed (99.9%) that a Senator was more right-wing than a Deputy of the same period. These effect sizes demonstrate that the findings of Desposato (2006), of no clear and consistent difference between the houses, cannot be supported.

Hypothesis 3 posits a difference in party behavior across the two institutions, with the literature paying particular reference to upper and lower house factions of the PMDB. Figure 3 plots the effect sizes of the differences in parties between the two institutions for the legislatures considered (the PT is plotted only in the 52nd legislature due to the low number of Senators the party had before this period). As can be readily observed from the graph, the effect size for the difference in the PMDB across the two houses is the more moderate of

²For each combination of means and standard deviations, the effect size is computed as $(\mu_1 - \mu_2) / \sqrt{((\sigma_1^2 + \sigma_2^2) / 2)}$. See Kruschke (2013).

³See Coe (2002) for interpretation of effect sizes.

Figure 1: Bayesian t-test of a difference in the means of the ideal points of the Chamber (μ_1) and the Senate (μ_2) for legislatures 48 to 53.

the three large parties, with this pattern changing only in the 52nd legislature. Among the four parties graphed, the PFL is the one that seems most split between the houses, at least until the 52nd legislature. Hence hypothesis 2 receives support for all major parties except the PMDB.

The ‘extremity’ argument of Bernabel, that senators within each party vote more for the party than deputies of the same party, is the inverse of Desposato’s claim that the parties in the Senate are less cohesive. One way to examine these arguments is to analyze the standard deviations of the t-test for a difference between the party means of both houses. Greater standard deviations translate to less cohesion, as the ideal points of the individuals in the party are more spread out in this case. The results are that there is exactly *one* instance of a party that more cohesive in the Senate than the Chamber – the PFL in the 52nd legislature, supporting Desposato’s findings and contrary to those of Bernabel. These results are shown in Table 1. Columns three and four show the spread of the 95% HDI for the posterior distributions of the standard deviations of the Chamber and Senate, respectively, and as can be seen from a comparison of these two columns, parties in the Senate have a wider HDI; they are therefore less cohesive.

Figure 2: Effect sizes for the difference in means between the Chamber and Senate, 48th – 53rd legislatures.

Party	Legislature	Chamber	Senate
PFL	48	0.1658342	0.3559811
PFL	49	0.1194475	0.3101234
PFL	50	0.1073254	0.2368030
PFL	51	0.1164046	0.2191900
PFL	52	0.2114796	0.1939089
PFL	53	0.4978111	0.5258737
PMDB	48	0.1528147	0.1832714
PMDB	49	0.1123227	0.2813139
PMDB	50	0.1658978	0.1687208
PMDB	51	0.1626198	0.2119743
PMDB	52	0.1392673	0.2429299
PMDB	53	0.1309777	0.3679892
PSDB	48	0.2445454	0.2865050
PSDB	49	0.1714863	0.3355876
PSDB	50	0.1263158	0.2743876
PSDB	51	0.1228523	0.3328872
PSDB	52	0.2480893	0.3839572
PSDB	53	0.1372616	0.3927237
PT	53	0.1646568	0.3107340

Table 1: Differences in the standard deviations of party means across the Chamber and Senate.

Figure 3: Effect sizes of the difference in mean ideal points for main parties in the Chamber and the Senate, legislatures 48 to 53. Zero is highlighted on the y-axis; effect sizes close to zero indicate a smaller difference between parties in the two houses.

5. Conclusion

I have shown in this brief research note article how the results of a formal comparison between two legislative institutions can differ greatly depending on the methodology employed in making the comparison. Indeed, the findings presented here are in many cases the opposite of those in other studies. For methods such as ideal-point estimation, where we are creating scales of measurement, it is imperative that positions on such scales are made properly comparable in order for us to avoid the possibility of making incorrect inferences. It is also worth noting that the points raised here apply as well to studies that compare subsets of the voting database to the whole, for example, in the comparison of foreign policy voting with voting on domestic policy. As ideal-point research on legislative institutions grows in Brazil and other countries in Latin America, it is important that comparisons are properly made. Of course, the substantive question of why these differences exist between upper and lower houses is an interesting avenue for future research. Using the methodology espoused here, and in [Bailey \(2007\)](#), [Treier \(2011\)](#) and [Bailey et al. \(2015\)](#), will ensure that the comparisons we make are as accurate as possible.

A. Appendix

Project Year	Chamber Date	Senate Date
MPV0025/89	1989-01-26	1989-01-26
MPV0032/89	1989-01-27	1989-01-27
MPV0027/89	1989-02-14	1989-02-14
MPV0028/89	1989-02-14	1989-02-14
MPV0031/89	1989-02-14	1989-02-14
MPV0035/89	1989-02-14	1989-02-14
MPV0159/90	1990-04-04	1990-04-04
MPV0151/90	1990-04-08	1990-04-08
MPV0154/90	1990-04-09	1990-04-09
MPV0168/90	1990-04-11	1990-04-11
MPV0288/90	1991-01-10	1991-01-10
MPV0291/91	1991-01-23	1991-01-23
MPV0014/01	2002-04-10	2002-04-16
MPV0107/03	2003-04-29	2003-05-21
MPV0114/03	2003-05-20	2003-06-11
MPV0116/03	2003-05-22	2003-06-12
MPV0118/03	2003-06-03	2003-06-24
MPV0120/03	2003-07-29	2003-08-12
MPV0135/03	2003-12-18	2003-12-22
MPV0136/03	2004-01-21	2004-02-05
MPV0144/03	2004-01-28	2004-03-09
MPV0147/03	2004-02-03	2004-03-17
MPV0168/04	2004-03-23	2004-05-05
MPV0167/04	2004-05-04	2004-05-19
MPV0179/04	2004-05-19	2004-06-08
MPV0182/04	2004-06-02	2004-06-17
MPV0225/04	2005-02-23	2005-03-08
MPV0241/05	2005-06-01	2005-06-28
MPV0248/05	2005-06-22	2005-08-10
MPV0255/05	2005-10-04	2005-10-26
MPV0281/06	2006-04-26	2006-06-07
MPV0316/06	2006-11-21	2006-12-06
MPV0336/06	2007-02-26	2007-03-27
MPV0360/07	2007-05-15	2007-06-27
MPV0392/07	2007-11-06	2007-11-27
MPV0402/07	2008-03-05	2008-04-09
MPV0408/07	2008-04-02	2008-05-06
MPV0409/07	2008-04-08	2008-05-06
MPV0424/08	2008-05-28	2008-07-09

Table 2: MPV votes used as bridging votes in the analysis.

References

- Bailey, M. A. (2007). Comparable Preference Estimates across Time and Institutions for the Court, Congress, and Presidency. *American Journal of Political Science*, 51(3):433–448.
- Bailey, M. A., Strezhnev, A., and Voeten, E. (2015). Estimating Dynamic State Preferences from United Nations Voting Data. *Journal of Conflict Resolution*, pages 1–27.
- Bernabel, R. (2015). Does the Electoral Rule Matter for Political Polarization? The Case of Brazilian Legislative Chambers. *Brazilian Political Science Review*, 9(2):81–108.
- Clinton, J. D., Jackman, S., and Rivers, D. (2004). The statistical analysis of roll call data. *American Political Science Review*, 98(02):355–370.
- Coe, R. (2002). It’s the effect size, stupid: What effect size is and why it is important.
- de Freitas, A. M., Izumi, M., and Medeiros, D. B. (2012). O Congresso Nacional em Duas Dimensões: Estimando Pontos Ideais de Deputados e Senadores (1988-2010). Paper presented at the 8th Meeting of ABCP 01-04 August, Gramado-RS.
- Desposato, S. W. (2006). The Impact of Electoral Rules on Legislative Parties: Lessons from the Brazilian Senate and Chamber of Deputies. *Journal of Politics*, 68(4):1018–1030.
- Izumi, M. Y. (2011). Governo e oposição no senado brasileiro. *bolsa*, 3153:7.
- Jackman, S. (2001). Multidimensional analysis of roll call data via bayesian simulation: Identification, estimation, inference, and model checking. *Political Analysis*, 9(3):227–241.
- Kruschke, J. K. (2013). Bayesian Estimation Supersedes the t Test. *Journal of Experimental Psychology: General*, 142(2):573–603.
- Kruschke, J. K. and Meredith, M. (2015). BEST: Bayesian Estimation Supersedes the t-Test. <https://CRAN.R-project.org/package=BEST>.
- Martin, A. D. and Quinn, K. M. (2002). Dynamic Ideal Point Estimation via Markov chain Monte Carlo for the US Supreme Court, 1953–1999. *Political Analysis*, 10(2):134–153.
- Neiva, P. and Izumi, M. (2012). Os sem-voto do legislativo brasileiro: quem são os senadores suplentes e quais os seus impactos sobre o processo legislativo. *Opinião Pública*, 18(1):1–21.
- Neiva, P. and Soares, M. M. (2013). The Brazilian senate: a federative or a party politics chamber? *Revista Brasileira de Ciências Sociais*, 28(81):97–115.
- Neiva, P. R. P. (2008). Os poderes dos senados de países presidencialistas e o caso do brasil. *O SENADO FEDERAL BRASILEIRO NO PÓS-CONSTITUINTE*, page 23.
- R Core Team (2015). R: A language and environment for statistical computing. <https://www.R-project.org/>.
- Shor, B., Berry, C., and McCarty, N. (2010). A bridge to somewhere: Mapping state and congressional ideology on a cross-institutional common space. *Legislative Studies Quarterly*, 35(3):417–448.

- Treier, S. (2011). Comparing Ideal Points Across Institutions and Time. *American Politics Research*, 39(5):804–831.
- Zucco Jr., C. and Lauderdale, B. E. (2011). Distinguishing Between Influences on Brazilian Legislative Behavior: Brazilian Legislative Behavior. *Legislative Studies Quarterly*, 36(3):363–396.

Affiliation:

Robert Myles McDonnell
Researcher, Fundação Getúlio Vargas
São Paulo, Brazil
E-mail: robertmylesmcdonnell@gmail.com
URL: <http://robertmyles.github.io/>

SocArXiv on Twitter

Temporary Home of SocArXiv Preprints

Preprint

osf.io/view/socarxiv/

<https://twitter.com/socarxiv>

<https://osf.io/view/socarxiv/>

Submitted: September 30, 2016

Accepted: September 30, 2016