Political Analysis (2003) 11:275–288 DOI: 10.1093/pan/mpg005

Comparing Group and Subgroup Cohesion Scores: A Nonparametric Method with an Application to Brazil

Scott W. Desposato

Department of Political Science, 315 Social Sciences, University of Arizona, Tucson, AZ 85710 e-mail: swd@u.arizona.edu

This article builds a nonparametric method for inference from roll-call cohesion scores. Cohesion scores have been a staple of legislative studies since the publication of Rice's 1924 thesis. Unfortunately, little effort has been dedicated to understanding their statistical properties or relating them to existing models of legislative behavior. I show how a common use of cohesion scores, testing for distinct voting blocs, is severely biased toward Type I error, practically guaranteeing significant findings even when the null hypothesis is correct. I offer a nonparametric method—permutation analysis—that solves the bias problem and provides for simple and intuitive inference. I demonstrate with an examination of roll-call voting data from the Brazilian National Congress.

1 Introduction

Roll-call cohesion scores are a staple of legislative scholarship. They have been in use since at least the 1920s (Rice 1924), and are still regularly used with only slight modifications.¹ They are flexible, intuitive, and simple to calculate. Interpretation is easy and widely understood. As a result, scholars have used these scores to study legislative roll-call votes in extremely diverse settings, including the Confederate Congress, the Russian Duma, the Argentinean Chamber of Deputies, and of course the Congress of the United States.²

Perhaps because the scores are simple, intuitive, and flexible, there has been little attention to understanding their statistical properties or how to apply them to models of legislators' roll-call votes. Cohesion scores are generated by the behavior of individual legislators, and there is a massive literature on individual legislators' voting decisions. But scholars typically shy away from any attempt to relate cohesion scores to individual legislators' voting calculus, relegating any such theorizing to a statistical black box. Instead of beginning with a theoretically driven data-generating function, most works treat cohesion scores as raw data and assume they have "nice" statistical properties. For inference, the usual approach is simply to use *t* tests applied to simple regression models or comparisons of means.

Author's note: Thanks to Ben Bishin, Barbara Geddes, Jeff Gill, Mark Jones, Kris Kanthak, Brian Lawson, and two anonymous reviewers. Special thanks to Moema Bonelli for excellent research assistance.

¹Rice (1925) notes, however, that other authors were using a similar cohesion index prior to his work.

²See Haspel et al. (1998), Jones (2000), Key (1949), and Schickler and Rich (1997).

Political Analysis, Vol. 11 No. 3, © Society for Political Methodology 2003; all rights reserved.

In this article, I show that one common use of cohesion scores is severely biased toward the alternate hypothesis. In some settings, this bias virtually guarantees finding significant effects where there are none. The problem is most evident in comparisons of group and subgroup cohesion scores, for example, floor vs. committee cohesion, as explained below. I suggest a nonparametric permutation test as an unbiased solution. This approach is robust and easy to implement; it also has the additional benefit of preserving the simple intuition that characterizes the widely used cohesion score.

The article proceeds in four additional steps. The next section reviews previous applied work and demonstrates the bias problem associated with this methodology. Section 3 provides a simple unbiased solution: permutation analysis. Section 4 applies this methodology to a current scholarly debate. Section 5 concludes.

2 Existing Research and Its Biases

2.1 Group vs. Subgroup Analyses

The bias problem I investigate arises when comparing group and subgroup average cohesion scores. What is the usual motivation for these comparisons? Typically, scholars are testing to see if there exists, in Rice's words, a "voting bloc." More specifically, scholars have asked, "do subgroups define divided coalitions within legislature?" The group–subgroup comparison framework encompasses a wide variety of legislative voting phenomena. For example, we might ask whether regional differences divide parties, perhaps testing whether Northern and Southern Democrats form distinct voting blocs. The same basic approach could be applied to comparisons of committee voting with floor voting, or of ethnic, religious, or gender caucus behavior.

This category of hypothesis test has been used in diverse contexts. Rice (1928) is the first to suggest the general approach, arguing that a comparison of overall legislature cohesion with subgroups is an appropriate test for voting coalitions. In his own work (Rice 1925), he compares the cohesion of several potential blocs in the New York State Assembly: the Assembly as a whole, Democrats, Republicans, Radicals, Progressives, and economic interests (farmers, laborers, and socialists). In each case he finds that the subgroups are more cohesive than broader groups that encompass them. For example, upstate and urban Democrats are more cohesive than Democrats overall. This leads him to identify regional parties as significant and distinct voting blocs within the state legislature. He also applies this approach to the U.S. Senate, again finding distinct subgroups.

Key (1949) used a similar framework, examining regional party blocs in the House and Senate. He compared Southern Democrats, non-Southern Democrats, all Democrats, and Republicans. His results suggest that Southern and non-Southern Democrats are more cohesive than Democrats overall, providing evidence of regional divisions in the Democratic party.

Truman (1956) examines state-party delegation cohesion in the House of Representatives. His analysis shows that voting agreement among state-party delegations is higher than overall parties, and Truman attributes this to the provision of voting cues through personal networks of legislators. Gile and Jones (1995) test for differences between Congressional Black Caucus (CBC) members and all Northern and Southern Democrats. They find that the CBC is in fact quite cohesive, and attribute this to racial solidarity and shared values of the African American legislators.

The same approaches have been applied to comparative studies of legislatures. Mainwaring and Liñán (1997) and Samuels (1996) have searched for evidence of "federal effects" in national legislatures by comparing party and state-party cohesion. They have

found that state-party delegations are more cohesive than national parties, and argue that a federalist form of government naturally creates state divisions in parties and detracts from the formation of a national policy agenda.

In each case, the basic formulation of the test can be framed as follows. First, the cohesion on a roll-call considering bill *j* is defined as

$$C_j = \frac{|Y_j - N_j|}{Y_j + N_j},\tag{1}$$

where Y_j is the number of "yes" votes and N_j is the number of "no" votes on bill *j*. C_j is thus the overall cohesion of the legislature on bill *j*. More commonly, scholars compute the cohesion score party by party:

$$C_{jk} = \frac{|Y_{jk} - N_{jk}|}{Y_{jk} + N_{jk}},$$
(2)

where Y_{jk} is the number of "yes" votes cast by members of party k and N_{jk} is the number of "no" votes cast by members of party k on bill j.

As legislative bodies typically vote on many pieces of legislation, scholars calculate party k's average cohesion score C_k across all n bill j as

$$C_{k} = \frac{\sum_{j=1}^{n} C_{jk} W_{jk}}{\sum_{j=1}^{n} W_{jk}}.$$
(3)

Finally, average cohesion across all parties can be calculated as

$$C_{\rm P} = \frac{\sum_{k=1}^{m} \sum_{j=1}^{n} C_{jk} W_{jk}}{\sum_{k=1}^{m} \sum_{j=1}^{n} W_{jk}},\tag{4}$$

where C_{jk} is the cohesion of party k on bill j, n is the total number of bills considered, m is the total number of parties, and W_{jk} is a weight for party k's votes on bill j.³ Thus, C_P can be thought of as a system-wide measure of party cohesion.

Subgroup cohesion C_{SP} is calculated in almost the same way. Let *l* index each mutually exclusive subgroup. For example, subgroups might be state delegations in each party (California Republicans in Congress, Nevada Republicans in Congress, etc.). Instead of party cohesion on vote *j*, we are interested in state-party cohesion:

$$C_{jkl} = \frac{|Y_{jkl} - N_{jkl}|}{Y_{jkl} + N_{jkl}}$$
(5)

where Y_{jkl} is the number of "yes" votes and N_{jkl} is the number of "no" votes cast by legislators from state *l* in party *k* on vote *j*. Let *o* be the number of subgroups of interest.

³Various weighting schemes have been used. One example is a simple threshold where $W_{jk} = 1$ if a minimum number of legislators oppose the measure and a minimum number of votes are case. Rice (1925), for example, throws out all unanimous votes and all votes where less than six members were in the minority. Other scholars only include votes where party majorities are in opposition. Weights may also be assigned according to each bill's importance or controversy. Scholars (e.g., Carey 2001) might weight by how divided the legislative parties were on bill *j*. Nevertheless, these weighting schemes have no impact on the basic problem described in this article.

Instead of overall party cohesion, we calculate overall state-party cohesion:

$$C_{\rm SP} = \frac{\sum_{k=1}^{m} \sum_{j=1}^{n} \sum_{l=1}^{o} C_{jkl} W_{jk} V_{jkl}}{\sum_{k=1}^{m} \sum_{j=1}^{n} \sum_{l=1}^{o} W_{jkl} V_{jkl}}$$
(6)

In this case, W_{jk} is again the weight assigned to party k's vote on bill j. The new term V_{jkl} is an additional weight term that is specific to the subgroup l of party k on vote j. Typically, V_{jkl} is a weight controlling for the number of legislators in each subgroup. Notweighting delegations would effectively be counting Alabama's 2 Democrats as equal to California's 32. There are contexts where not weighting by size would be appropriate, but when comparing group and subgroup averages, not weighting can transform any correlation between group size and voting into apparently significant differences between group and subgroup cohesion.

Finally, scholars have compared group and subgroup cohesion using standard *t* tests for differences of means, using the following null and alternative hypotheses:

• H_0 : Subgroups do not define voting coalitions. Subgroups' average cohesion scores (C_{sub}) are equal to those of the overall group (C_{group}) :

$$C_{\text{group}} = C_{\text{subgroup}}.$$
(7)

• *H*_A: Subgroups *do* define distinct voting blocs within groups. Their expected cohesion scores are greater than that of the overall group:

$$C_{\text{group}} < C_{\text{subgroup}}.$$
 (8)

This approach has a number of laudable features. The test is intuitive and easy to understand, the math is simple and uncomplicated, and the calculations and coding are easy to perform. As a result, these voting bloc tests have been a common application of cohesion scores. They were used in early roll-call cohesion analysis and are still being used today, and they have been applied to a wide variety of important and relevant political science questions in many countries and time periods. Unfortunately, as we will see below, they are all biased.

2.2 The Problem with Comparisons of Group and Subgroup Cohesion Scores

The method described above is always biased toward Type I error, that is, toward rejecting the null hypothesis when it is in fact true. The problem is that under a standard model of legislative voting, the expected value of each subgroup's cohesion score is *always* greater than the broader group's cohesion score under the null hypothesis. This characteristic can create the illusion of voting coalitions, even when there are none. This is illustrated in the following paragraphs.

Consider a case where we are studying the cohesiveness of state-party delegations in a national legislature. Scholars have suggested that in federal systems, state-level interests create divisions in national political parties (Mainwaring 1999; Morgenstern 2000; Carey and Reinhardt 2001; Geddes and Benton 1997). In this case, national parties are the primary group, and state-party delegations are the subgroup of interest. Scholars tackling this problem have sought evidence that average state-party delegation cohesion is greater than overall national-party cohesion. The key statistics to be compared are average party cohesion (C_P) and average state-party delegation cohesion (C_{SP}).

Party	5–0	4-	-1		3–2			2–3		1-	-4	0–5
State A	3–0	2–1	3–0	3–0	2–1	1–2	2–1	1–2	0–3	1–2	0–3	0–3
State B	2–0	2–0	1–1	0–2	1–1	2–0	0–2	1–1	2–0	0–2	1–1	0–2

Table 1 Group vs. subgroup cohesion: sample space of roll-call outcomes (yes-no)

Table 2 Group vs. subgroup cohesion: corresponding sample space of cohesion scores

Party $(C_{\rm P})$	1	.6	5		.2			.2			5	1
State A	1	.33	1	1	.33	.33	.33	.33	1	.33	1	1
State B	1	1	0	1	0	1	1	0	1	1	0	1
$C_{\rm SPW}$	1	.6	.6	1	.2	.6	.6	.2	1	.6	.6	1
$C_{\rm SP}$	1	.67	.5	1	.17	.67	.67	.17	1	.67	.5	1

I illustrate the bias inherent in these comparisons using a simple example: the cohesion score for a single party on a single roll-call vote. For simplicity, it is supposed that there are only two states, A and B, and the party has five legislators, three from state A and two from state B.

Tables 1 and 2 illustrates this example. Table 1 shows the universe of all possible vote outcomes (Yes–No) for the overall party and state-party delegations. Table 2 shows the cohesion scores that correspond to each possible vote outcome, including overall cohesion, cohesion for State A and State B, and average state cohesion. Two measures of average state-party cohesion are provided: C_{SPW} , which is weighted by state-party delegation size, and C_{SP} , which is unweighted.

Beginning with the weighted state-party mean C_{SPW} , note that state-party cohesion scores are *always* equal to or greater than national-party cohesion scores—never lower. Why is this? Whenever the majority of each state-party votes the same way as the overall national party, the state and national scores will be the same. For example, on the second hypothetical outcome, an overall majority voted "yes" (4–1), a majority of State A voted "yes" (2–1), and a majority of State B voted "yes" (2–0). As the subgroup majorities align with the overall group majority positions, the national and state cohesion scores (C_{P} and C_{SPW}) are the same: .6.

But when the majority of any state party's members vote contrary to the national party's majority position, the state weighted mean will be higher. For example, on the fourth outcome in Table 1, the overall majority voted "yes" (3–2), State A voted "yes" (3–0), and State B voted "no" (2–0). In this case, the overall cohesion score is .2 while mean state-party cohesion (C_{SPW}) is 1.0. This happens because whenever a state's majority goes against the national majority, this pulls down the overall national average more than the state average. The individual states can still look relatively cohesive while the national party looks divided.

The problem is that under most random-utility models of legislative voting, all of the vote outcomes are possible with some positive probability—even if there are no federalist effects. Using a simple example, let each member of the party have an equal probability p of voting "yes" on that bill, and probability 1 - p of voting against it. In this case, as each legislator's behavior on a vote follows a Bernoulli distribution, the overall party vote of k "yes" votes and n - k "no" votes follows a binomial distribution with parameters n = 5, k, and p.

Under this simple model, there are no state or subgroup effects. *Every* legislator's behavior is driven by the same underlying mechanism, and there are no differences between states.

Expected Values of Party and Party-State Cohesion Under the Null Hypothesis 1.0 0.8 Expected Cohesion 0.6 0.4 0.2 0.0 0.2 0.0 0.4 0.6 0.8 1.0 P(Vote Yes)

Fig. 1 Bias in cohesion scores' tests.

But in spite of this, we could randomly see state-party cohesion scores that are higher than national-party cohesion scores. Consequently, as state-party cohesion is always equal to or greater than national-party cohesion, the expected value of state parties will always be greater than that of national parties, except under perfect party cohesion when p = 0 or p = 1.

How do these results change when using unweighted state-party cohesion (C_{SP}) ? In this case, it is sometimes possible to observe state-party cohesion below that of overall party cohesion, but the subgroup bias persists anyway. For example, on a 3–2 split, with State A voting 2-1 and State B voting 1-1, unweighted average state cohesion is .17, below the overall average of .2. This happens whenever the smaller subgroups' votes are less cohesive than that of the larger subgroups.

Nevertheless, bias is measured through expected values, not individual outcomes. It turns out that the expected value of $C_{\rm SP}$ continues to be above the expected value of $C_{\rm P}$ on any particular vote outcome, with or without weights. Tables 1 and 2 provide some intuition as to why. While subgroup cohesion can sometimes be lower than group cohesion, these outcomes are balanced by other outcomes where subgroup cohesion is much higher. For example, on a 3–2 overall vote, C_{SP} can be 1.0, .17, or .67. There are 10 ways this vote could take place $\binom{5}{3}$, and under the simple model presented above, each of the 10 permutations is equally likely to occur. One outcome produces a C_{SP} of 1.0, six lead to a C_{SP} of .17, and three lead to a C_{SP} of .67. So the expected value of state-party cohesion on a 3–2 overall vote is

$$E(C_{\rm SP} \mid (3yes, 2no)) = \frac{1}{10} \times 1.0 + \frac{6}{10} \times .17 + \frac{3}{10} \times .67 = .4$$
(9)

which is greater than the overall cohesion of .2 (and equal to expected weighted state-party cohesion). A general proof that expected subgroup cohesion is greater than expected group cohesion for two or more subgroups and any positive weighting scheme is provided in the Appendix.

Figure 1 plots the expected values for national-party and state-party cohesion as a function of p, under the assumptions presented above.⁴ The solid line represents the expected national

280

Scott W. Desposato

⁴Each legislator has an independent and identical probability p of voting "yes"; there are two states, with three legislators from one state and two from the other.

22.29

MPG005

May 16, 2003

cohesion score $E(C_P)$, and the dotted line represents the expected state-party cohesion score $E(C_{SP})$. As argued, the expected value for state-party cohesion is *always* above national-party cohesion. The difference is greatest where cohesion is lowest (p = .5) and the two values converge at the extremes (p = 1 or p = 0).⁵

This shows that any group vs. subgroup comparison of cohesion scores can produce an apparently significant difference where there is none. Scholars are virtually guaranteed to find that subgroups are more cohesive than groups. I illustrated this phenomenon in a simple case of a single vote, single party, and just two state subgroups. But the same mechanism is at work with any larger number of parties, roll-call votes, or subgroups. The extent of the expected bias will vary with several parameters.

First, as seen in Fig. 1, the lower the overall party cohesion, the greater the observed difference. Cohesion is lowest where the probability of voting "yes" is equal to the probability of voting "no," i.e., p = .5. At that point we observe the largest difference between state-party and national-party cohesion. At the extreme values, where the probability of voting yes is equal to 1 or 0, there is no difference between group or subgroup scores.

Second, the pattern is most likely to emerge in small state delegations or small subgroups. The smaller the delegation, the more likely that a majority of its members will vote the "wrong" way just by chance. For example, in a state with just three members and with p = .8, this will happen about 10% of the time.⁶ In a state with 20 members, the probability of this event is less than .001.⁷

3 A Solution

This does not mean that previous works' findings are wrong—just that all previous measures are biased toward a false positive finding. In this section, I propose and implement a nonparametric permutation test that avoids the bias problem associated with previous methods.

Permutation tests require minimal assumptions, but allow us to explore the distribution of the statistic of interest under the null hypothesis. The test uses a simple logic to build a distribution of the statistic of interest under the null hypothesis. First, assume that there are no subgroup effects—that every legislator in the larger group (in the example above, the party) casts a vote that is an independent draw from the same underlying distribution. Note that no assumptions about the form that distribution takes are necessary.

With these basic assumptions, we can see that any reordering of a roll-call vote is equally possible. For example, with five legislators, as above, and a final vote of 4–1 in the party, each of the outcomes in Table 3 is an equally likely outcome under the null hypothesis.

⁵Why is expected cohesion greater than zero when p(yes) = .5? The easy explanation is that there is only one way to get E(cohesion) = 0, which is a perfect 50–50 split in a party, but there are many ways to get cohesion scores above zero. Any vote outcome other than a 50–50 split will produce a cohesion score above zero. So the expected value when p(yes) = .5 has to be greater than zero. Desposato (2002) shows that the extent to which cohesion is inflated above zero is a function of party size—largest for small parties, and converging toward zero as party size goes to infinity.

⁶Binom (p = .8, n = 3, k = 1) + Binom (p = .8, n = 3, k = 0) = .104.

⁷Subgroup cohesion can artifically appear to be lower than overall group cohesion when there are many state-party delegations of size 1. If there were many states with only one member of party Z, and all voted "yes," and there were several states with two members of party Z who split their votes, state-party cohesion would be lower than national party cohesion. This only happens, however, because the single-member states fall out of the cohesion calculations at the state-party level. By definition, we need more than one member per state-party to calculate a state-party cohesion score. The solution is to drop observations that form subgroups of size 1 before calculating both overall and subgroup cohesion.

	Votes with 4–1 outcomes						
Deputy	1	2	3	4	5		
A	Yes	Yes	Yes	Yes	No		
В	Yes	Yes	Yes	No	Yes		
С	Yes	Yes	No	Yes	Yes		
D	Yes	No	Yes	Yes	Yes		
E	No	Yes	Yes	Yes	Yes		

 Table 3
 A simple example of permutations

Note. Each vote is equally likely under H_0 .

Hence exploring the distribution of the statistics of interest (C_N and C_{SP}) under various reorderings (permutations) provides insight into their distributions under the null hypothesis. We then compare these permutated values with the actual observed values. Practically, the test is implemented as follows.

- 1. Within each national party-vote combination, randomly reorder all the votes cast as a new variable. In other words, on each bill, randomly shuffle the Democrats' votes, and redeal them to that party. Then do the same for the Republican party.
- 2. Calculate the state-party cohesion scores from this new data set, as described above. Save the result.
- 3. Repeat this process many times, saving all the results.

The resulting density is an approximation to the distribution of the statistic of interest under the null hypothesis. That is, the saved party cohesion values are what we would expect to see if there really were no federalist impact on party cohesion, and if legislators' votes were not affected by state-based considerations.

We can then compare the actual value observed using the original dataset and ask does it fall within the range of normally occurring values under the null hypothesis? If the observed value is greater than 95% of the permutated values, we can say that the achieved significance level, or ASL, is .05. If there are federalist effects, the observed value should be above the distribution of permutated values.⁸

4 An Application

I illustrate the bias in traditional methods and apply a permutation test to one practical example. Scholars in recent years have used group vs. subgroup comparisons to explore the impact of federal forms of government on legislative party systems. The basic arguments are as follows. Federal forms of government create and/or strengthen state-level (or provincelevel) interests. Legislators in many national Congresses are accountable to state-level actors for nomination, election, or career advancement. Consequently, national parties may split on questions that mobilize pressure or that greatly affect state interests. To the extent that these forces are at work, national-party cohesion should be lower because of the federal form of government.

One country where this argument has been applied is Brazil. A federal republic with 26 states and a federal district, Brazilian state politics have been characterized as having

282

May 16, 2003

22.29

MPG005

⁸See Effron and Tibshirani (1994) for more details on the permutation test and other nonparametric statistics.

Table 4	Federalism in the Chamber of
Dep	uties mean cohesion scores

	National party	State party
1991–1994	.78 (.004)	.84 (.004)
1995–1999	.78 (.003)	.82 (.002)

Note. Standard errors are in parentheses.

Table 5	Federalism in the Senate mean
	cohesion scores

	National party	State party
1991–1994	.67 (.009)	.84 (.013)
1995–1999	.68 (.004)	.82 (.007)

Note. Standard errors are in parentheses.

important national effects. Brazilian federalism places key political resources in the hands of state actors, granting them influence over the behavior of national legislators and weakening national-party leaders (Mainwaring 1997, 1999; Selcher 1998; Souza 1998; Ames 2001; Samuels 2002). For example, initial nominations to run for Congress and access to free media time for campaign advertisements are both controlled by state parties. Similarly, state governors distribute many political jobs, including coveted directorships of state agencies. Consequently, "Politicians of the catchall parties focus a lot on state and local issues, so they are less likely to toe the line of the national party leadership" (Mainwaring 1997, p. 83). As a result, "...[f]ederalism influences the party system because most key decisions are made at the state level and abundant resources are allocated at this level" (Mainwaring 1999, p. 263).

A comparison of group vs. subgroup average cohesion scores supports these authors' claims. Table 4 shows national-party and state-party delegation average cohesion scores for the Chamber (lower house) and Senate for the last two legislative sessions.⁹ In each case, the state party cohesion scores are on average significantly higher than those of the national party. These results support the idea that state-party delegations do form voting blocs that weaken national party cohesion. The differences in cohesion scores are not huge, but are all significant at the .01 level using a simple *t* test.¹⁰

Table 5 presents the same data for the Senate. As with the Chamber, state parties are always more cohesive than the national parties, regardless of measure or period. The differences here are larger—over .10 in several cases. Again, the evidence would suggest that federalist forms of government can weaken national political parties.

⁹In this case, I weighted each party-bill and party-state-bill cohesion score by the number of legislators in each subgroup and by the overall cohesion of the vote. Specifically, V_{jkl} is the number of legislators in state *l* from party *k* that cast votes on roll-call *j*. W_{jk} in this case is the overall cohesion level of the legislature (constant for all parties on each vote), or $\frac{|Y_j - N_j|}{Y_j + N_j}$. As explained above, the trend produced by bias is identical using other weighting schemes. Using Rice's approach (Rice 1925) of just throwing out bills without minimal opposition, cohesion scores would all look higher. Using Carey's approach (Carey 2001), all cohesion scores would look lower. But the basic pattern of higher state party cohesion than national party cohesion would persist regardless of weighting scheme.

¹⁰Data come directly from the Brazilian Chamber and Senate. Coding and sources are discussed in Desposato (2001).



Fig. 2 State-party delegation size, 1995.

However, as I explained above, a positive and significant difference is virtually guaranteed by the bias inherent in these comparisons. In this case, we do not know if the results reflect real political differences or are simply a "false federalism" induced by the use of biased statistics.

In fact, it happens that Brazil is ideally suited to give a false positive to researchers using cohesion scores to test for federalist effects. As discussed above, the bias in these kinds of comparisons is most pronounced in settings with low cohesion and small subgroups. Brazil's political arena has both of these characteristics. Party cohesion is in fact not very high, and there are many states with small delegations to the National Congress.

Figure 2 shows the distribution of state-party delegation size in Brazil's Chamber of Deputies, based on the 1994 Congressional election results. Clearly, the vast majority of state-party delegations are quite small. Fully 44% are of size 1, another 20% are of size 2, and 90% are of size 6 or less. There are a few delegations of more than 10 members, but these only represent about 4% of all delegations.

All these small delegations have a much higher probability of randomly voting against the overall party majority—without any federalist effects. The large number of small state-party delegations virtually guarantees that scholars will find a federalist impact when comparing party and state-party cohesion scores even if there is none. The pattern is only exacerbated in the Brazilian Senate. Each state elects three Senators, and most state-party delegations are of size 1, and the largest are only of size 3. The minute state-party delegations of the Senate magnify the false federalism bias of cohesion scores, and explain why federalist effects appear stronger in the Senate than in the House.

Figure 3 applies the permutation test described above to the Brazilian case. The four graphs hold results for the Chamber and Senate in each of two periods, 1991–1994, and 1995–1998. The histogram shows the frequency of state-party cohesion based on 1000 random permutations of the original dataset. The solid line shows the observed national-party cohesion ($C_{\rm P}$). The dashed line shows the observed average state-party cohesion ($C_{\rm SP}$).¹¹

¹¹Note that the empirical values of party cohesion from Fig. 3 (the solid vertical lines) differ from the party cohesion values in Tables 4 and 5, because state-party delegations of size 1 have been dropped from the calculation of national party cohesion for the permutation analysis. The standard approach, presented in the table, compares group and subgroup average cohesion without any such adjustments.



Fig. 3 1000 permutations of state-party cohesion scores.

Several patterns stand out. First, the figures all clearly show the bias in state-party cohesion scores. Even where there were no federalism effects, *all* the 1000 permutated state-party cohesion scores were well above the national-party score. As I suggested, Brazil is an ideal case for producing false positives in this particular test.

Second, even so, there *are* consistent federalist effects observable in the data. The observed state-party cohesion scores are well above what we would expect to observe under the null hypothesis. While the impact of federalism on party cohesion is only about half as large of what we previously observed, this does provide strong evidence that state interests do divide national political parties. These results are consistent for both the Senate and Chamber of Deputies, for both periods.

5 Conclusion

Roll-call cohesion scores are a staple of legislative scholars. They have been widely used across country, time, and institution. Their calculation is simple and intuitive, they avoid complicated modeling of individual legislators' behavior, and are easy to program. An important application of cohesion scores has been to the identification of distinct voting blocs within legislatures. In particular, scholars have compared group and subgroup cohesion scores.

In this article, I demonstrated that such comparisons produce conclusions biased toward Type I error. Specifically, subgroup cohesion scores are virtually *guaranteed* to be more higher than overall group scores. Where parties are highly cohesive and subgroups are large, the bias in group vs. subgroup cohesion comparisons is small. But when subgroups are small or cohesion low, the bias is large and Type I error virtually assured.

I proposed a nonparametric permutation analysis to correct for this "false federalism" bias. The method preserves the best characteristics of the Rice cohesion score: it is simple,

easy, and intuitive. When applied to a test case of the Brazilian legislature, the results illustrate the severe bias in simple cohesion score analyses. The results also confirm that there are significant state-party cohort effects on Congressional voting behavior in Brazil, though the effects are quite small. The approach could be applied to a variety of legislative research problems, including the study of ethnic, religous, and gender blocs within legislatures, as well as regional blocs within parties.

Appendix: Proof of Group-Subgroup Difference

This appendix shows that the expected value of subgroup cohesion scores is greater than the expected value of overall group cohesion scores. Let C_G be the Rice cohesion score for the overall group, C_{Si} be the Rice cohesion score of subgroup *i*, where all subgroups are mutually exclusive, of size 2 or more, and all legislators are part of a subgroup, and let \overline{C}_S be the average Rice cohesion score for all subgroups. Demonstrating the bias requires showing that

$$E(\bar{C}_{\rm S} - C_{\rm G}) > 0.$$
 (A1)

I begin with a model of legislative voting. Let p be the probability of voting "yes" and 1 - p be the probability of voting "no." Let n be the number of legislators in the overall group. Assume there is only a single roll-call vote. Under this simple model, each legislator's behavior follows a Bernoulli distribution with probability p, and the overall vote outcome follows a binomial distribution with probability p, number of trials n, and number of successes ("yes" votes) k. The expected Rice cohesion score for the overall group is

$$E(C_{\rm G}) = \sum_{k=0}^{n} \operatorname{Rice}(k) P(k).$$
(A2)

Q1

Desposato (2002) showed that the expected value of a cohesion score could be reduced as follows.

$$E(C_{\rm G}) = \sum_{k=0}^{n} \frac{\|k - (n-k)\|}{n} \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}.$$
 (A3)

Rewriting to eliminate the absolute value signs, if n is even,¹² produces

$$E(C_{\rm G}) = \sum_{k=0}^{n} \frac{2k-n}{n} \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} + 2\sum_{k=0}^{\frac{n}{2}-1} \frac{n-2k}{n} \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$$
(A4)

and can be reduced to

$$E(C_{\rm G}) = (2p-1) + 2\sum_{k=0}^{\frac{n}{2}-1} \frac{n-2k}{n} \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}.$$
 (A5)

¹²Generalizing to *n* odd only requires replacing the $\frac{n}{2} - 1$ term with $\frac{n-1}{2}$.

Defining

$$B(n) = 2\sum_{k=0}^{\frac{n}{2}-1} \frac{n-2k}{n} \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$$
(A6)

gives us

$$E(G) = (2p - 1) + B(n).$$
(A7)

The second term, B(n), can be thought of as the impact of a majority of the party voting the "wrong" way—voting "no," when p(yes) is greater than .5. Desposato (2002) shows that as party size or p increases, this term goes to zero because the probability of a majority voting the wrong way goes to zero. With a party of three and p(yes) of .6, the probability of a majority (2 or 3 of 3) voting "no" is about .35; with a party of 200, the probability of 101 or more voting "no" is effectively zero. The substantive result, discussed by Desposato (2002), is that small groups are guaranteed to look more cohesive than larger groups—even if they are not.

This result makes a demonstration of the group–subgroup bias straightforward. Let there be *n* legislators in the group, each having an independent probability *p* of voting "yes" and 1 - p of voting "no" on a bill. Let there be *m* mutually exclusive subgroups of at least two members within the larger group, let every member of the larger group belong to exactly one subgroup, and let n_i denote the size of the *i*th subgroup.

We can now write

$$E(\bar{C}_{\rm S} - C_{\rm G}) \tag{A8}$$

$$= E(\bar{C}_{\rm S}) - E(C_{\rm G}) \tag{A9}$$

$$= E(\bar{C}_{S}) - (2p - 1) - B(n).$$
(A10)

As $E(\bar{C}_{\rm S})$ is the expectation of a sum, we write

$$E(\bar{C}_{S}) - (2p-1) - B(n) = \sum_{i=0}^{m} \frac{E(C_{Si})}{m} - (2p-1) - B(n)$$
(A11)

$$= \sum_{i=0}^{m} \frac{(2p-1) + B(n_i)}{m} - (2p-1) - B(n)$$
 (A12)

$$= (2p-1) + \sum_{i=0}^{m} \frac{(B(n_i))}{m} - (2p-1) - B(n) \quad (A13)$$

$$=\sum_{i=0}^{m} \frac{(B(n_i))}{m} - B(n).$$
 (A14)

As all n_i are smaller than n, all of the $B(n_i)$ are larger than B(n). As all of the $B(n_i)$ are larger than B(n), their mean is also larger, independent of any positive weighting or lack thereof. Consequently,

$$E(\bar{C}_{\rm S} - C_{\rm G}) > 0,$$
 (A15)

as we set out to demonstrate.

Q2

Scott W. Desposato

References

288

Ames, Barry. 2001. The Dadlock of Democracy in Brazil. Ann Arbor: The University of Michigan Press.

- Carey, John M. 2001. "Party and Coalition Unity in Legislative Voting." Paper presented at the American Political Science Association Meetings.
- Carey, John M., and Gina Yannitell Reinhardt. 2001. "Coalition Brokers or Breakers: Brazilian Governors and Legislative Voting". Unpublished.
- Desposato, Scott W. 2001. Institutional Theories, Societal Relatities, and Party Politics in Brazil. Ph.D. Thesis. Los Angeles: University of California.
- Desposato, Scott W. 2002. "Correcting for Bias in Roll-Call Cohesion Scores." Working Paper.

Effron, Bradley, and Robert J. Tibshirani. 1993. *An Introduction to the Bootstrap* (Monographs on Statistics and Applied Probability, No. 57). New York: Chapman and Hall.

Geddes, Barbara, and Allyson Benton. 1997. "Federalism and Party Systems." Unpublished.

Gile, Roxanne L., and Charles E. Jones. 1995. "Congressional Racial Solidarity: Exploring Congressional Black Caucus Voting Cohesion, 1971–1990." *Journal of Black Studies* 25:622–641.

Haspel, Moshe, Thomas F. Remington, and Steven S. Smith. 1998. "Electoral Institutions and Party Cohesion in the Russian Duma." *Journal of Politics* 60:417–439.

Jones, Mark P. 2000. "Explaining the High Level of Party Discipline in the Argentine Congress." Unpublished. Key Jr., V. O. 1949. *Southern Politics in State and Nation*. New York: Alfred A. Knopf.

- Mainwaring, Scott. 1997. "Mutipartism, Robust Federalism, and Presidentialism in Brazil." In *Presidentialism and Democracy in Latin America*, eds. Scott Mainwaring and Matthew Soberg Shugart. New York: Cambridge University Press, pp. 55–109.
- Mainwaring, Scott P. 1999. *Rethinking Party Systems in the Third Wave of Democratization: The Case of Brazil.* Stanford, CA: Stanford University Press.
- Mainwaring, Scott, and Aníbal Pérez Liñán. 1997. "Party Discipline in the Brazilian Constitutional Congress." Legislative Studies Quarterly 12:453–483.

Morgenstern, Scott. 2000. "Explaining Voting Unity in the Legislatures of the United States and Latin America." Paper Presented at the annual meeting of the American Political Science Association.

Rice, Stewart A. 1925. "The Behavior of Legislative Groups: A Method of Measurement." *Political Science Quarterly* 40:60–72.

Rice, Stewart A. 1928. Quantitative Methods in Politics. New York: Alfred A. Knopf.

Rice, Stuart A. 1924. Farmers and Workers in American Politics. Ph.D. thesis. Columbia University.

Samuels, David J. 1996. "Legislative Lilliputians?' Towards a Theory of Party Cohesion in the Brazilian Chamber of Deputies." Unpublished.

- Samuels, David J. 2002. "Ambassadors of the States: Political Ambition, Federalism, and Legislative Politics in Brazil." Unpublished.
- Schickler, Eric, and Andrew Rich. 1997. "Controlling the Floor: Parties as Procedural Coalitions in the House." American Journal of Political Science 41:1340—1375.
- Selcher, Wayne. 1998. "The Politics of Decentralized Federalism, National Diversification, and Regionalism in Brazil." *Journal of Interamerican Studies and World Affairs* 40:25–50.

Truman, David. 1956. "The State Delegations and the Structure of Party Voting in the United States House of Representatives." *The American Political Science Review* 50:1023–1045.

Sourza, Celina. 1998. "intermediação de Interesses Regionais no Brasil: Impacto do Federalismo e da Descentralização." *Dados* 41:569–592.

P1: *** MPG005 May 16, 2003 22:29

Queries

Q1. Au: Should this "P" be in lowercase [i.e. p(k)]?

Q2. Au: Please confirm spelling of first name "Stuart".