How we count counts: The empirical effects of using coalitional potential to measure the effective number of parties

Reuben Kline*

University of California, Department of Political Science, 3151 Social Science Plaza, Irvine, CA 92697, USA

Article info

Article history:
Received 20 August 2008
Received in revised form 24 November 2008
Accepted 7 January 2009

Keywords:
Effective number of parties
Voting power indexes
Party systems

Abstract

Despite its conceptual centrality to research in comparative politics and the fact that a single measure—the Laakso–Taagepera index ($N_s$)—is widely employed in empirical research, the question of what is the best way to “count” parties is still an open one. Among other alleged shortcomings, $N_s$ has been criticized for over-weighting small parties, especially in the case of a one-party majority. Using seat shares data from over 300 elections, I have for the first time calculated $N_s$ as well as an alternative measure ($N_{Bz}$) which employs Banzhaf scores, rather than seat shares, as weights. The Banzhaf index is a voting power index which calculates a party’s voting power as a function of its coalitional potential. Though the two measures are highly correlated, I identify three particular party constellations in which the differences between $N_s$ and $N_{Bz}$ are significant and systematic.

1. Introduction: electoral success versus governance

In comparative political research, the need to quantify the number of parties operating within a political system is fundamental [see, for example, Sartori (1976), Lijphart (1994, 1999), Taagepera (2007), Taagepera and Shugart (1989) among many others]. Despite its conceptual centrality to research in comparative politics and the fact that a single measure—the Laakso–Taagepera index—is extensively, if not exclusively, employed in comparative research, the question of what is the best way to “count” parties is far from obvious. This paper will argue that the dominant approach—one which is based on using either party seat shares or party vote shares as the weighting factor in the construction of the index—in general provides a more accurate measure of the number of parties which are competitive (based on the preceding election) at a given time, or over time. On the other hand, one can apply many other weighting factors to the general Laakso–Taagepera framework (what will be called the reciprocal self-weighted average below). A measure to be further elaborated below, $N_{Bz}$, employs normalized Banzhaf-adjusted voting power scores as weighting factors. Because these weighting factors represent the share of all possible winning coalitions in which a party is pivotal from a coalition building standpoint, this index can provide a more intuitive measure of the parties which have a potential for governing after any given election.

This index—one member, of many, of a class of indices which measure voting power—was developed by John Banzhaf (1965) as a way to measure the relative power of a voter in an assembly.1 In the application we are concerned with here, the Banzhaf index assigns weights to parties as a function of the relative frequency that each, when considering the set of all possible winning coalitions, is a “swing” voter.

Thus, $N_{Bz}$ incorporates a certain conception of coalitional viability into the party weighting scheme. As a result, given certain party configurations, the two indices, $N_s$ and $N_{Bz}$, can give strikingly different results. Though the indices

---

1 In fact, the index was originally proposed by Penrose (1946), in the Journal of the Royal Statistical Society. Banzhaf’s (1965) publication appeared in the Rutgers Law Review. It was not until Coleman’s (1971) treatment of the same index, that the index became widely known in the social sciences.
are indeed highly correlated, below I identify three types of party constellations in which the differences between them are systematic and statistically significant.

After a brief review of the methods previously proposed for counting parties in the next section, I discuss the most recent attempts to amend or replace $N_e$, including the use of normalized Banzhaf scores as party weights. Using data from 329 post-World War II elections spanning 24 countries, I then examine the differences in the two distributions and identify three cases in which the two indices produce systematically divergent values. This paper represents the first attempt to systematically compare empirical results based on these two important measures of the effective number of parties ($ENP$).

Finally, though it has often been the case when new indexes are proposed, I do not attempt to demonstrate the generic superiority of this new Banzhaf-based measure. Rather, I argue that it is better to view it simply as one member of a class of indexes of a certain format—the reciprocal self-weighted average (RSWA) format—a class in which the members are differentiated solely by their weighting factors. When conceptualized in this way it becomes easier to view the indexes as complements rather than substitutes. Moreover, given that the same underlying formula is used to construct all of these measures, we can be confident that any differences between two measures of the RSWA format are due entirely to differences between the weighting factors. As a result, comparing and contrasting these various indexes (for example, by using their differences or ratios) becomes obviously useful.

2. A brief accounting for the way we count

Duverger (1959), in a seminal study which laid the foundations for his eponymous “law” regarding the effect of the electoral system on the number of parties, merely counted the parties that were in competition for seats. While this crude approach has simplicity to recommend it, it became clear that it was necessary to somehow weight each party in order to give a more accurate measure of the effective number of parties for comparative purposes.

Blondel (1968) undertakes such this task. He develops a typology of two, two-and-a-half, and multiparty systems. There are essentially two problems with such an approach. First, for many purposes a continuous measure of parties is needed. Second, the cutoff point for “half” and “strong” parties is essentially arbitrary (he uses approximately 10% and 40% respectively).

While not precisely a measure of the number of parties, Rae’s (1971) fractionalization index was the first attempt to construct a measure which is continuous and takes into account all parties which have won seats, while also systematically weighting them by their seat shares. Rae’s formula is worth reproducing here:

$$F = 1 - \sum s_i^2$$

(1)

Where $s_i$ is the proportion of legislative seats for party $i$. The index, based on seat shares of every party in the system, is a useful summary of the relative size and number of parties in a system. Nonetheless, its interpretation is not straightforward, and it does not give a ready measure of the number of parties operating in a system.

Because of this, scholars continued to work on a simple yet more readily interpretable single measure to represent the shape of a political party system. Laakso and Taagepera (1979) develop what would become the standard for comparative political research. Construction of their $N_s$ index involves only the same (minimal) amount of data that is required for the fractionalization index. It is calculated as follows:

$$N_s = \frac{1}{\sum_{i=1}^n (s_i)^2}$$

(2)

Where, again, $s_i$ is the seat share of party $i$ in the parliament.

It is also worth pointing out its relationship to Rae’s fractionalization index:

$$N_e = \frac{1}{\sum_{i=1}^n (s_i)^2} = \frac{1}{1 - F}$$

(3)

Though seat shares are used in Eq. (2), the effective number of parties is often calculated using party-vote shares as well. Moreover, many different weighting factors for the individual parties can be applied to this general format. One specific weight, the normalized Banzhaf index, will be discussed in the next section, and the possibility of creating compound measures consisting of some combination of two or more $ENP$ constructs will be discussed further below. All $ENP$ measures based on the RSWA format will be designated by a sub-scripted $N$, where the subscript will denote which weighting factor is being used. For example, $N_d$ refers to the $ENP$ measure with seat shares as the weighting factor, whereas $N_v$ would refer to the index based on vote share weighting, and $N_BZ$ will refer to the index which employs (normalized) Banzhaf scores as party weights. Another version of the $ENP$ index introduced in Blau (2008)—one which bases its party weighting factors on the cabinet portfolio (ministerial) shares of each party—to be denoted here as $N_{cab}$, will be discussed further below.

2.1. Laakso–Taagepera: current debates

According to Arend Lijphart, “in modern comparative politics a high degree of consensus has been reached on how exactly the number of parties should be measured” (Lijphart, 1994, p. 68). Despite this claim regarding the nearly universal agreement on a disciplinary standard, there have been several attempts to elaborate substitutes or complements to $N_e$. These include Molinar (1991), Taagepera (1999), Dunleavy and Boucek (2003), and Dumont and Caulier (2003).
Recognizing that in the case of absolute dominance (i.e., a single-party majority), \(N_s\) can sometimes produce seemingly unrealistic values, Taagepera (1999) proposes a supplementary indicator—the reciprocal of the largest party’s seat share—in an attempt to obviate this irregularity. This auxiliary measure, despite having the appealing property of falling below two only in the case of absolute dominance, is nonetheless only supplementary and cannot be used on its own.

Molinar (1991) combined \(N_s\) with the largest party seat share, to create an index denoted NP. While the values which obtain under NP are generally quite similar to those of the largest component approach, NP can yield a value less than two even when there are perhaps three or more parties which are relevant in the coalition building sense (Taagepera, 1999).

Dunleavy and Boucek (2003) consider an index, Nb, which is simply the average of \(N_s\) and the largest party seat share. They claim that this “produces a highly correlated measure, but one with lower maximum scores, less quirky patterning and a reader interpretation.” Though Nb does yield results which are marginally more intuitive in some cases than those produced by \(N_s\), it fails to entirely correct the problem created by a single-party majority party system.

While \(N_s\) remains the most frequently used measure of party system shape, there are certain seat share distributions in which it is likely to produce misleading results. Many of the disadvantages of \(N_s\) (as well as \(N_p\)) stem from the fact that it tends to overestimate the actual number of relevant parties, by giving excess weight to parties which are entirely irrelevant from the standpoint of governance in any given election. The index weights each party by its (proportional) seat share, but does so without regard to the distribution of the remainder of the seats among the other parties. For example, if a given party holds 10% of the seats, its weight in the index will be the same whether the remaining 90% of the seats are evenly divided between nine other equally-sized parties or are all occupied by one super-dominant party. In other words, the index does not take into account the coalition building potential of each party. Though here the term coalition will be used to refer to a governing coalition, the logic of the Banzhaf index can be more or less isomorphically applied to the case of ad-hoc coalitions formed to pass specific legislation.

It is widely held that party competition tactics are strongly influenced by the number of parties. Sartori (1976, p. 120) writes, “... in particular, the tactics of party competition and opposition appear related to the number of parties; and this has, in turn, an important bearing on how governmental coalitions are formed and are able to perform.” From this quote it is clear that Sartori envisions competition as taking place in two distinct arenas: the electoral arena and the legislative arena. These two aspects of competition manifest themselves in electoral competition, and coalition formation (i.e., post-election bargaining within the legislature) respectively. Moreover, there is a feedback between the number of parties and various aspects of coalition government. This means that the number of parties is important as both a dependent and independent variable in many models of politics, rendering a proper understanding of its measurement all the more important.

In the decades following Sartori’s seminal piece, this distinction would continue to be discussed. Laver (1989, p. 301) writes, “The process of party competition is generally divided, by both theorists and empirical researchers, into a number of component parts. Two of the most important of these components are electoral competition and legislative behavior.” For the purposes of this exposition, the most beneficial aspect of \(N_{Bz}\) is that it incorporates both of these crucial aspects of competition identified by Laver. On the one hand, \(N_{Bz}\), like \(N_s\), takes into account seat shares—which are the ultimate result of electoral competition. Crucially, \(N_{Bz}\) goes one step further by also incorporating simplified, yet nonetheless useful, aspects of the “politics of coalition.” Moreover, \(N_{Bz}\) incorporates this second important feature of party competition without requiring any more data than what is required by \(N_s\).

According to Dumont and Caulier (2005), \(N_s\) is usually “interpreted in comparative political science as the number of hypothetical equal-sized parties competing or being influential for the building of a majority government.” It would appear that when we are talking about “competing” in this sense, it must mean electoral competition. Thus, it would make sense if we have an application which requires the use of the effective number of parties in a context that is meant to reflect electoral competition—e.g., electoral volatility—then the \(N_s\) is likely to be a more useful measure. However, if we encounter an application in which it is important to consider the effective number of parties from the standpoint of coalitional viability—e.g., cabinet duration—then \(N_{Bz}\) may be the wiser choice.

Because of the anomalous behavior of the \(N_s\) in several special types of party constellations (more on that below), an alternative construction of the party weights has been suggested by several scholars, including—most recently and elaborately—Dumont and Caulier (2005). This way of constructing the party weights ensures that the parties are

---

4 Schofield (1993) also notes the dual nature of party competition. He puts forth a model where “parties are concerned with policy outcomes but choose party positions both with a view to electoral consequences and as a basis for coalition bargaining.” Here, Schofield is concerned with policies. The reason policy has not been addressed in this paper is because the Banzhaf index is an a priori power index, meaning that it treats the formation of all possible winning coalitions as equally likely. While this may not be ideal from a predictive or case-study standpoint, it has decided advantages for large-N cross-national comparative research. For criticism of the Banzhaf index in this vein see Gelman et al. (2002) and Margolis (1983).

5 Though the basic data required is the same, due to the necessity of identifying all possible winning coalitions for each distribution of seat shares, the computational requirements are substantially greater for \(N_{Bz}\). Despite this, there are two websites, which given a seat share distribution and a decision rule, will automatically calculate voting power indexes (including Banzhaf, Shapley-Shubik, Coleman, Zipf, and Owen among others). One website, called Powerslave, is hosted by the University of Turku’s (Finland) research group on voting power: http://powerslave.val.utu.fi/ Another is maintained by Dennis and Robert Leech of Warwick University and the University of London respectively: http://www.warwick.ac.uk/~ecaae/. In this project, the Powerslave application was used to calculate all of the BZ scores.

6 Independently, Grofman (2006) briefly discusses the possibility of constructing such a measure.
weighted according to their potential to form a part of the governing coalition. Therefore, parties with different vote shares but nonetheless identical coalitional potential (as specifically defined here) will have identical weights. This aspect of \( N_{Bz} \) is more amenable to identification of certain broad types of party configurations, as a large set of different seats-share distributions can lead to identical \( N_{Bz} \) values.\(^7\) This fact will become clearer in Section 4.2.

Notationally, \( N_{Bz} \) and \( N_s \) differ only in their subscripts. This difference reflects the sole difference in their construction. They are indices of the same format, differing only in their respective party weighting factors. This format, the RSWA mentioned above differs, however, from some of the other formats that have been suggested, such as an entropy-based format due to Wildgen (1971).

Before discussing the Banzhaf index in more detail in the next section, a few other important points regarding the use of the RSWA format are in order. First, the RSWA is unconcerned with the type of weight used. Thus, the format, not the weighting factor, is what defines this class of index. The only restriction on the weights is that they must sum to one. Otherwise, there are infinite possibilities for weights (though, in practice, there are certainly a finite, and relatively small number of weights which would make intuitive sense) which can be plugged into this basic format. In addition to ENP measures based on votes and seats (\( N_v \) and \( N_s \) in our notation), Blau (2008) introduces two new ENP measures: one based on legislative power and one based on cabinet power. Each of these uses the RSWA format, though he makes explicit that his notion of legislative power which is based on Powell (2000) is distinct—because it includes more precise and context-specific qualitative information—from the notion of legislative power based on any form of voting power index.

These works demonstrate that the RSWA format can be used efficaciously for ENP measures based on votes, seats or cabinet shares, as well as more difficult to measure concepts such as ‘legislative power’ (as in Blau, 2008). Importantly, the self-weighted nature of this format ensures that, regardless of the conceptual basis of the weights which are used, the contributions of less “weighty” parties are heavily discounted,\(^8\) and any observed difference among indices of this type are purely a result of the different conceptual basis employed. If we were to consider in turn \( N_v, N_s, N_{Bz}, \) and \( N_{cab} \), we should observe, in general, \( N_v > N_s > N_{Bz} > N_{cab} \). I do not wish to belabor this particular point here, but it is important to note that the recognition of RSWA as a single format can allow for useful comparisons of various RSWA-based ENP measures. In this case, looking at the differences among the indices as we move from \( N_v \) to \( N_{cab} \) should allow us to make some inferences regarding the competitiveness of the party system in question at each of these stages of the electoral and governance processes.

2.2. The Banzhaf Index

The Banzhaf Index is just one of a class of a priori voting power indices. Others include, \textit{inter alia}, Shapley–Shubik, Coleman and Owen. Voting power indices can be used in any case where \textit{blocks} of votes exist, and it is reasonable to assume that the blocks—at least in terms of voting—are unitary actors, i.e., they vote as a block.\(^9\) A voting power index can be applied in any case where there is \textit{weighted} voting. An a priori index requires only two inputs: the decision rule and the distribution of vote shares. In our case, the decision rule is simple majority (\( 50\% + 1 \) votes), and the distribution of vote shares is merely the proportion of seats controlled by each party in the parliament.

In this construction, each party is weighted by counting the number of times it is a \textit{swing} voter out of all possible winning (i.e. majority) coalitions (WCs). This number is then normalized by dividing by the total number of swings out of all of the possible winning coalitions. A party is defined as a swing within the context of a particular (winning) coalition if its removal from a coalition renders an otherwise winning coalition a losing one.

The normalized Banzhaf score is calculated by taking the number of times party \( i \) is a swing voter divided by the aggregate number of swings in all possible winning coalitions. For party \( i \) the index would be:

\[
\frac{Sw_i}{\sum_{i=1}^{n} Sw_i} = Bz_i
\]

Where \( Sw_i \) is the number of times \( i \) is a swing voter.

The final step is analogous to the construction of the effective number of parties (\( N_e \)):

\[
N_{Bz} = \frac{1}{\sum_{i=1}^{n} (Bz_i)^2}
\]

Where \( Bz_i \) is the normalized Banzhaf score for party \( i \), replacing \( s_i \) as the weight.

This index, \( N_{Bz} \), referred to by Dumont and Caulier (2005) as the Effective Number of Relevant Parties (ENRP) can avoid many of the anomalies displayed by \( N_e \), while introducing others. The addition of “relevant” is an explicit reference to Sartori’s notion of relevance; i.e., a party having either coalitional or blackmail potential. In the construction of the Banzhaf weights, both of these aspects of coalition building are taken into account. A given party’s number of swings is indicative of both its utility as a coalitional partner and the leverage it has by threatening to defect from the coalition. Both of these “potentials” are ultimately derived from the nature of the Banzhaf index itself, which measures the relative propensity for a party to be a swing player in a coalition, and thus takes into account

---

\(^7\) While such a situation is theoretically possible in the case of \( N_v \), it is highly improbable.

\(^8\) I want to thank Rein Taagepera especially for enlightening my thinking with regard to the RSWA index and the way in which it could potentially be used to measure the ENP concept (and many other things!).

\(^9\) Thus the Banzhaf index can be used—and in fact was originally constructed to measure—legislative apportionment. Additional uses include shareholder voting and weighted voting in international bodies such as the European Union.
the power of each party based both on its membership in all possible winning coalitions, as well as its ability to threaten to turn the winning coalition into a losing one.

In the case of one-party dominance, \( N_{Bz} \) always gives a value of one (because the only swing voter is the party with the majority of seats), thus avoiding the sometimes counterintuitive result displayed in the usage of \( N_s \). Moreover, \( N_{Bz} \) can be calculated using the very same data necessary for construction of \( N_s \). On the other hand, \( N_{Bz} \) is more computationally intensive, and in some cases the index can take on the same value based on widely disparate party configurations.\(^{10}\) Keeping in mind that each index has its advantages and disadvantages, in the next section I identify and examine three types of party configurations in which the two indices yield systematically divergent values.

3. An empirical comparison of \( N_{Bz} \) and \( N_s \)

3.1. Comparison of the two distributions

Using data from Mackie and Rose (1997) I calculate \( N_s \) and \( N_{Bz} \) for 329 post-World War II elections in 24 countries. Table 1 provides summary statistics for these calculations. As might be expected I find that the mean is lower for \( N_{Bz} \), partly a result of the fact that, as discussed above, it always returns a value of 1 when there is a single-party majority. Also, note that standard deviation is higher.

The difference in the means is highly significant \((p < 0.01)\).

Table 2 displays the (arithmetic) mean values of each of the indices by country as well as the difference between two index values. Note that in all cases except two, the indices by country as well as the difference between two index values is more or less linear. With this scatter plot, the tendency for low values of the two indices, the relationship appears to be highly correlated, the more interesting cases are those in which they differ dramatically and/or systematically.

Table 2

<table>
<thead>
<tr>
<th>Country</th>
<th>( N_s )</th>
<th>( N_{Bz} )</th>
<th>( N_s - N_{Bz} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>2.50</td>
<td>1.92</td>
<td>0.58</td>
</tr>
<tr>
<td>Austria</td>
<td>2.45</td>
<td>2.34</td>
<td>0.12</td>
</tr>
<tr>
<td>Belgium</td>
<td>5.08</td>
<td>4.72</td>
<td>0.36</td>
</tr>
<tr>
<td>Canada</td>
<td>2.37</td>
<td>1.74</td>
<td>0.64</td>
</tr>
<tr>
<td>Denmark</td>
<td>4.57</td>
<td>3.75</td>
<td>0.82</td>
</tr>
<tr>
<td>Finland</td>
<td>5.03</td>
<td>4.75</td>
<td>0.28</td>
</tr>
<tr>
<td>France</td>
<td>3.41</td>
<td>2.27</td>
<td>1.14</td>
</tr>
<tr>
<td>Germany</td>
<td>3.18</td>
<td>3.46</td>
<td>0.28</td>
</tr>
<tr>
<td>Greece</td>
<td>2.19</td>
<td>1.66</td>
<td>0.53</td>
</tr>
<tr>
<td>Iceland</td>
<td>3.84</td>
<td>3.38</td>
<td>0.46</td>
</tr>
<tr>
<td>Ireland</td>
<td>2.82</td>
<td>2.30</td>
<td>0.52</td>
</tr>
<tr>
<td>Israel</td>
<td>4.44</td>
<td>4.50</td>
<td>0.05</td>
</tr>
<tr>
<td>Italy</td>
<td>3.79</td>
<td>3.44</td>
<td>0.35</td>
</tr>
<tr>
<td>Japan</td>
<td>2.94</td>
<td>1.53</td>
<td>1.41</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>3.30</td>
<td>3.14</td>
<td>0.15</td>
</tr>
<tr>
<td>Malta</td>
<td>2.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Netherlands</td>
<td>4.67</td>
<td>4.60</td>
<td>0.07</td>
</tr>
<tr>
<td>New Zealand</td>
<td>1.96</td>
<td>1.05</td>
<td>0.91</td>
</tr>
<tr>
<td>Norway</td>
<td>3.35</td>
<td>2.30</td>
<td>1.05</td>
</tr>
<tr>
<td>Portugal</td>
<td>3.01</td>
<td>2.03</td>
<td>0.98</td>
</tr>
<tr>
<td>Spain</td>
<td>2.71</td>
<td>1.93</td>
<td>0.78</td>
</tr>
<tr>
<td>Sweden</td>
<td>3.33</td>
<td>2.40</td>
<td>0.93</td>
</tr>
<tr>
<td>Switzerland</td>
<td>5.21</td>
<td>4.46</td>
<td>0.74</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>2.12</td>
<td>1.30</td>
<td>0.83</td>
</tr>
</tbody>
</table>

\(^{10}\) On this last point I would especially like to thank Filippo Tronconi.

ranges from about 2 to about 4. Given their frequencies and the systematicity with which they appear, these clusters will be investigated more thoroughly below.

Although, in general, \( N_{Bz} \) is more likely to produce lower values than \( N_s \), when a country has many parties (like the Netherlands) the difference is slight. However, as is discussed below, there are three particular types of distributions which are likely to produce widely disparate values. These are the case of one-party majority, and the two different cases of particular three-party constellations. However, since super-dominant majorities are rare, then the most realistic one-party majorities are those which are barely majorities—those that are just slightly above 50%. With bare majorities, however, comes an increased likelihood of party seat shares just under 50%. Since two of the three cases where the two measures are most likely to drastically diverge are likely to be present across time in the same electoral environment, it could be the case that the two effects may cancel each other out, and then on average, the effective number of parties may be fairly similar. To summarize, this implies that by using averages across time, we may be under-estimating the true magnitude of the difference between the two measures. Perhaps this would be an indication that \( N_s \) might be better interpreted as the shape of the party system over time.

First, in order to be sure that we are dealing with statistically different distributions, some sort of non-parametric test is required. I employ the Wilcoxon rank-sum test, a test in which the null hypothesis is that the two
distributions are identical. Using the Wilcoxon test, we obtain $z = 4.73$, which leads to the rejection of the null hypothesis (i.e., the claim that the two distributions are identical) at $\alpha = 0.01$. Therefore, while highly correlated, it is clear that these two indices are measuring different things. It is the goal for the remainder of this project to explore these differences.

As a first attempt to analyze the systematic differences in the distributions, consider Fig. 2.

Fig. 2 displays the cumulative frequency distributions of $N_s$ and $N_{Bz}$. Notice that the case of $N_s$, the values of the index are slowly but steadily increasing, reflecting the fact that even very small changes in seat shares will cause a change in the index value. In the case of $N_{Bz}$, however, we can note two vertical segments in its cumulative distribution function. These occur at $N_{Bz} = 1$ and $N_{Bz} = 3$. There are 92 cases of $N_{Bz} = 1$, meaning that in 28% of our cases, a single-party controls the majority of the parliamentary seats, as such a scenario is the only one which produces a $N_{Bz}$ value of exactly one. Thus, absolute dominance is not only theoretically interesting from our point of view, but empirically important as well.

The second plateau occurs at $N_{Bz} = 3$. In 18.5% of our cases, $N_{Bz}$ is exactly equal to 3, which can occur if and only if exactly three parties share all of the possible swings equally among them. In this case, however, there are two different types of party configurations which will yield this index value. The two configurations are one in which there are three roughly equally sized parties with many smaller parties (all with absolutely no coalitional potential), and one in which there are two roughly equally sized large parties (with seat shares in the 40% range) and a third, but much small party, which is large enough to form a winning coalition with either of the other two parties.

In these three types of party constellations, $N_{Bz}$ and $N_s$ differ systematically. In the next section, I examine more in depth these three types of party configurations.

3.2. Three illustrative cases

As we have seen, $N_s$ generally yields large values than $N_{Bz}$. In fact, Dunleavy and Boucek (2003) and Molinar (1991) claim that $N_s$ exaggerates fragmentation by overstating the number of relevant parties. I will discuss three cases in which $N_s$ and $N_{Bz}$ differ significantly and systematically. In two of these cases, $N_s$ does in fact always produce lower values than $N_{Bz}$ but in the third, $N_{Bz}$ is quite often larger than $N_s$.

3.2.1. Constellation 1: single-party majority

In this case, in which one-party single-handedly has a parliamentary majority, we might expect the number of effective parties to be close to one, as one-party presumably has the power to single-handedly push legislation through the parliament. $N_{Bz}$, by construction, takes the value of one if and only if there is a one-party majority. In the case of $N_s$, however, the value under a majority party configuration can be as high as 3 or more. Given its empirical prevalence, then it follows that the method of counting political parties in this case can lead to significantly different classifications of party systems.

Given that nearly 30% of our cases are of a single-party majority, this case is empirically relevant. For an illustration
of the stark differences that obtain under this constellation, I use the example of Japan. In the post-World War II period, Japan has been regarded by many as a one-party democracy, with the Liberal Democratic Party (LDP) dominating the political landscape (both electorally and legislatively) until very recently. Our data include elections from 1946 to 1990, and show that, on average, \( N_s = 2.94 \) while \( N_{Bz} = 1.53 \). This high value of \( N_s \) is despite the fact that in 11 of the 18 elections held during this period the LDP obtained an absolute majority. In this case, \( N_s \) is on average nearly twice as high as \( N_{Bz} \). Fig. 3 displays the \( N_{Bz} \) and \( N_s \) values for each of these 18 Japanese elections under consideration. In addition to these important differences in Japan, in the 92 cases in which there is a one-party majority, the \( N_s \) is larger on average by 1.26 parties.

3.2.2. Constellation 2: balanced Tri-Partism

In a case where we have three medium, roughly equally-sized parties (each with the same number, two, of swings) and many small, non-pivotal ones, \( N_s \) is likely to be significantly larger than \( N_{Bz} \). Though this type of party configuration is relatively rare (at least in the data analyzed herein), it is an example of the most common critique of \( N_s \)—that it gives excessive weight to (irrelevant) small parties, thus exaggerating the extent to which the party system is fractionalized. I take for an illustrative example the results from the 1982 elections in the Netherlands, which are displayed in Table 3.

For this particular seat share distribution, \( N_s = 4.02 \) and \( N_{Bz} = 3. \) \( N_s \) is greater than 4 despite the fact that only three parties (those which are highlighted) have any coalitional potential. Which index is better to use in this situation would largely depend on the context in which it is being applied.

3.2.3. Constellation 3: Unequal Tri-Partism

Though it is typically the case the only party \( N_s \) is larger than \( N_{Bz} \), this constellation—what I have termed Unequal Tri-Partism—will always produce \( N_{Bz} \) values larger than the \( N_s \) values. This case is characterized by two large, roughly equally sized parties (with seat shares in the 40% range) and one smaller party that is able to form a winning coalition with either of the other two parties.\(^{11}\) The reason that in this case \( N_{Bz} \) is higher than \( N_s \) is that \( N_s \) is not giving as much weight to the third, smaller party. Though \( N_{Bz} \) will always be higher in this case, the difference between the two will be less to the extent that there are more (and larger) non-pivotal parties. The size of the non-pivotal parties, taken together, is of course constrained by the combined size of the three relevant parties.

The two Austrian examples found in Table 4 are typical for this constellation. Moreover, they are typical for the country as well. In fact, nine out of the 16 Austrian elections in the sample result in this constellation. Always in such a situation \( N_{Bz} = 3 \), and in the Austrian cases of this phenomenon, the \( N_s \) value ranged from 2.1 to 2.9. However, this constellation manifests itself in other countries as well. In total, there are 36 cases of this in the data, with the \( N_{Bz} \) being greater, on average, by 0.43 parties.

4. Discussion

While several scholars have advocated the wholesale replacement of \( N_s \), it nonetheless remains the measure that is used in nearly all applications. This is not without good reason. \( N_s \) provides an easily calculated and readily interpretable unique measure of party system shape. Attempts at modifying the underlying formula to correct for some of \( N_s \)'s alleged shortcomings, such as Molinar (1991) and Dunleavy and Boucek (2003), have been largely unsuccessful. While they may partially ameliorate some of \( N_s \)'s limitations, they come with their own set of deficiencies, and are, in most cases, more difficult to construct.

\( N_{Bz} \) (or ENRP to use Dumont and Caulier's terminology) is an attempt to merely redefine the weights attributed to each party. These redefined weights are based on the coalitional potential of each party rather than simply on their seat shares. In redefining the weights thusly, the measure takes on correlated, yet in some cases strikingly divergent,

\(^{11}\) A further necessary condition, noted above, is that these three parties are the only with any swings.
values. Given the high degree of correlation between the two measures, for many party configurations the choice between these two indices may not be terribly important. However, in the three particular configurations (Single-Party Majority, Balanced Tri-Partism, and Unequal Tri-Partism) identified above, the choice may be quite important, and the ultimate choice should depend on the context in which the measure is applied. If one is interested in a phenomenon, such as electoral volatility, which is inherently defined over time, then $N_s$ would seem to be a more fitting measure. If, on the other hand, one is interested in an election-specific phenomenon, such as cabinet duration, then $N_{Bz}$ will perhaps give a more accurate representation of the true state of affairs.

Table 3
Balanced Tri-Partism in the 1982 Dutch elections.

<table>
<thead>
<tr>
<th>Party</th>
<th>Seat share (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labor Party</td>
<td>31.3</td>
</tr>
<tr>
<td>Christian Democratic Appeal</td>
<td>30.0</td>
</tr>
<tr>
<td>People's Party for Freedom and Democracy</td>
<td>24.0</td>
</tr>
<tr>
<td>Democrats '66</td>
<td>4.0</td>
</tr>
<tr>
<td>Pacifist Socialist Party</td>
<td>2.0</td>
</tr>
<tr>
<td>Political Reformed Party</td>
<td>2.0</td>
</tr>
<tr>
<td>Communist Party of the Netherlands</td>
<td>2.0</td>
</tr>
<tr>
<td>Political Party Radicals</td>
<td>1.3</td>
</tr>
<tr>
<td>Reformatory Political Federation</td>
<td>1.3</td>
</tr>
<tr>
<td>Reformed Political Alliance</td>
<td>0.7</td>
</tr>
<tr>
<td>Center Party</td>
<td>0.7</td>
</tr>
<tr>
<td>Evangelical People's Party</td>
<td>0.7</td>
</tr>
<tr>
<td>Ns</td>
<td>3.0</td>
</tr>
<tr>
<td>Ns</td>
<td>4.0</td>
</tr>
</tbody>
</table>

Though it may be difficult to delineate precisely the contexts in which any single measure of the effective number of parties is unequivocally superior, it is clear that using some combination of them—perhaps a difference or a ratio—can be useful. This is especially the case if one wishes to consider the institutional, cultural and historical effects on the party system as a whole. For example, few would argue that $N_v = N_s$ or $N_v/N_s$ are not reasonable measures for the disproportionality of an electoral system. Similarly, one could consider $N_{Bz} - N_{cab}$ or $N_{Bz}/N_{cab}$ as measures of the extent to which coalitional power is translated into power over policy portfolios. In such a way, analogous measures could be constructed to examine the extent to which norms, institutions and history conspire to nearly inevitably reduce the effective number of parties as we consider each subsequent stage of the electoral and governance process. Perhaps merely analyzing these differences between various RSWA measures employing different weighting factors can provide some insight into the extent to which party systems restrict party and policy competition at each of these stages. One additional conceptual extension of the RSWA format of ENP measures is the consideration of ideological dispersion and/or policy distance as party weighting factors. Having such a wide array of potential ENP measures could allow us to tackle such questions as how disproportionality and polarization relate to one another, or enable us to analyze the effect of a priori voting power on the ideological composition of governing coalitions.

Finally, despite the usefulness of combinations of various ENP measures, particular research questions may nonetheless require the use of a single ENP measure. As demonstrated by the high degree of correlation between $N_{s}$ and $N_{Bz}$ noted above, determining which measure is optimal, even for a given context, is a difficult task. Hopefully this work will serve as a useful springboard for such efforts.

Acknowledgements

The author would like to acknowledge the helpful comments of Bernie Grofman, Rein Taagepera, Don Saari Marek Kaminski and Filippo Tronconi and two anonymous referees. Earlier drafts of this paper were presented at the 2007 SWPSA meetings, and the Center for the Study of Democracy Graduate Student Conference at UC Irvine in 2008. In addition, the author is grateful for financial support from the Institute for Mathematical Behavioral Sciences and the Center for the Study of Democracy, both at UC Irvine. Any remaining errors and omission are, of course, the sole responsibility of the author.

References


