4

The Number of Parties and Proportionality – Two Key Tools for Analysis

The number of parties is among the most frequent numbers in political analysis, and it is central to the study of party systems. Similarly, the "proportionality" of an electoral system is one of the most frequently discussed criteria for evaluating a system. A party system involves, of course, much more than the mere number of parties, but it is impossible to describe it without giving some idea of how many players are involved. Similarly, an electoral system can be assessed on many criteria other than proportionality. Nonetheless, these two tools for analysis – the number of parties and proportionality – are so central to the analysis of electoral systems and party systems that we devote this entire chapter to these two concepts.

THE NUMBER OF PARTIES

Characterizing the types of party systems has concerned students of party politics for a long time – see the excellent overview by Wolinetz (2006). What is a meaningful number of parties in an assembly? As we discussed already in Chapter 1, just counting parties is not always useful. Some parties are large and others small, and thus we need some way to have a count that emphasizes the number of parties that are "important" in some sense.

In this book, we will use the index known as the *effective number of parties*. This index was first proposed by Laakso and Taagepera (1979). It has become the industry standard (e.g., Lijphart 1994; Cox 1997) even as various others have been proposed and used by some scholars.[1] The effective number is a size-weighted count, devised to give more weight to the largest party and less to the smaller ones. It may be calculated on either seat shares or vote shares. The formula for calculating the effective number of seat-winning parties (which we can designate N_S) is:

[1] For examples, see Molinar (1991), Dunleavy and Boucek (2003), and Golosov (2009).

$$N_S = \mathbf{1}/\sum(s_i)^2 = \textbf{inverse sum of squared fractional shares}.$$

In words, we square the seat shares for each of *i* parties – however many there are, starting with the largest party, s_1. Then we sum up all the squares. Once we have this sum, we take the reciprocal. In this way, the index weights each party by its size. The squaring results in a large party contributing more to the final index value than does a small one. For instance, suppose the largest party has half the seats, $s_1 = 0.5$; thus we have $0.5^2=0.25$. Now suppose among several remaining parties a smallest (the i^{th}) one has only 5 percent of the seats. We take the share, 0.05, and square it, and get 0.0025. In this way, when we sum up the squared shares of all the parties, the smallest one has counted for much less than the largest. This is precisely what we want – a size-weighted count of how many parties there are. See Taagepera (2007: 47–64) for a more systematic and detailed treatment.

Alternatively, we could calculate our index on vote shares, giving us the effective number of vote-earning parties (N_V):

$$N_V = \mathbf{1}/\sum(v_i)^\mathbf{2}.$$

Here v_i stands for the fractional *vote* share of the i^{th} party. Thus for any given election result, we have two effective numbers: N_S for the seats and N_V for the votes.[2] These numbers are sometimes referred to as ENPP (Effective Number of Parliamentary Parties)[3] and ENEP (Effective Number of Electoral Parties), respectively. However, given our interest in systematically constructing logical models, we adopt the approach more typical of scientific notation: single symbols with subscripts.[4]

Table 4.1 offers a brief demonstration of how we get from vote or seat shares to the effective number of parties, using the above formula. It shows ten hypothetical examples of party systems, each consisting of two to five parties. The entries under the columns for each hypothetical parties could be either votes or seats shares. We square each share in any of these rows, sum of the squares, and then take the reciprocal of the sums. Then we have the effective number.

If we have some number of equal-sized parties, then the effective number is the same as the actual number – as it logically should be! So, in example no. 1, we have two parties, each with a share of 0.50, and we get N=2.00. Similarly,

[2] In addition, the effective number of components can be useful outside the realm of electoral and party studies, whenever well-defined components add up to a well-defined total. For instance, one can measure the effective number of polities in the world, based on their areas or populations (Taagepera 1997). One can measure the effective number of states or provinces in a federal system, and compare it to the actual number (Taylor et al. 2014). That is, the fractional shares that enter the equation for *N* can be shares of votes, seats, or anything else that we are interested in summarizing.

[3] Or sometimes Effective Number of Legislative Parties (ENLP).

[4] Moreover, ENP risks being mistaken for multiplication of the quantities *E, N* and *P*.

TABLE 4.1 *Examples of hypothetical party systems and resulting values for an effective number of parties*

	Votes (or seats) shares					
Example	Party 1	Party 2	Party 3	Party 4	Party 5	Effective N
1	**0.50**	**0.50**				**2.00**
2	0.667	0.333				1.80
3	0.667	0.167	0.166			2.00
4	0.50	0.25	0.25			2.67
5	**0.333**	**0.333**	**0.333**			**3.00**
6	0.50	0.33	0.17			2.58
7	0.333	0.333	0.25	0.084		3.43
8	**0.25**	**0.25**	**0.25**	**0.25**		**4.00**
9	0.25	0.25	0.25	0.13	0.12	4.57
10	**0.20**	**0.20**	**0.20**	**0.20**	**0.20**	**5.00**

with the other cases of equal sized parties in examples 5, 8, and 10, shown in bold.[5] In the second example, we still have only two parties, but one is substantially bigger than the other, yielding N=1.80. This illustrates a general pattern: if there are two unequal components, N<2. Then in example 3, the largest party stays the same size as in example 2, but there are now three parties. What should N do? It must go up! There are now three parties. However, because they are unequal, N could not be as high as the actual number. It must be in between. The formula yields N=2.00, intuitively making sense that it is between one (we have a hegemonic party) and three (we are pretty far from three equal-sized parties). In other words, N=2.00 is not always a two-party system! In most actual cases a value near 2.00 is likely to be closer to what we might mean by the vague "two-party system" than our made-up example.

The other examples in Table 4.1 illustrate the same basic points – we see N decrease when the largest party becomes larger for the same number of parties (compare example 6 to 5), but it goes up if the largest party stays the same but one of the other parties splits (compare example 7 to 5, or example 9 to both 8 and 10).[6] The effective number is never larger than the total number of parties receiving seats or votes.

[5] The reverse need not be the case. Thus N=2.00 might result from a balanced .50–50 (example 1), or from a very unbalanced .667–.167–.166 (example 3).

[6] At the same effective number, the parties' sizes may or may not be balanced. Compare for instance 34-33-33 and 53-15-10-10-10-2, both of which have N=3.00. Taagepera (2007: 50–53) offers a formula for an index of balance (which would result in 0.98 and 0.35, respectively, for these examples).

The formula for the effective number is "operational" in that it can be applied mechanically to any constellation of fractional shares. But the formula does not tell us which shares to use. Should the German CDU and its Bavarian "sister party" the CSU be counted as a single party or two? Either decision is justified, as the parties act as if they were one in some respects but retain their separate organizations and act independently in some other ways. The formula is not able to tell us which to do, only how to calculate once we decide whether to enter CDU and CSU shares separately or as a combined party. In the opposite direction, internal factions of a party, such as the Liberal Democratic Party in Japan, could be seen as "parties within the party" (Reed and Bolland 1999).

Lijphart (1999: 69–74) settles such dilemmas by calculating N both ways and taking the arithmetic mean. This is a reasonable solution when the two approaches are fairly close. But some cases present extra challenges. Consider the case of Chile's open-list PR system, where there are alliances, each consisting of several distinct parties. The effective number of *alliances* is close to two. However, the effective number of distinct *parties* is much higher (see our discussions of alliances in list-PR systems in Chapters 6 and 14). Both numbers make sense, in different ways, but their mean may make no sense at all. Our general practice in this book will be to count as a "party" those entities that *call themselves parties* and for which separate votes totals are reported. However, there are times when counting alliances rather than their component parties makes theoretical sense – for instance, when the alliance is essentially permanent across several elections and consistent in its membership across the country. Such is the case in Chile; India likewise has many parties that cooperate in alliances (see Chapters 5 and 15). Different researchers may make different decisions about how to count parties, alliances, and factions when calculating the effective number. The key is being consistent and transparent about what is being put into the formula.

Electoral Systems and Effective Number of Parties

It is widely understood that the electoral system somehow affects the number of parties. In particular, do multiseat PR systems produce more parties than does FPTP, in line with Duverger's (1954) propositions? It also has been claimed that presidential systems reduce the number of parties.[7] In **Table 4.2** we compare values of both N_V and N_S for FPTP and PR systems, and also for countries with parliamentary and presidential executives. We see that the ranges of N_V and N_S are definitely lower for FPTP than they are for multiseat PR. However, these are overly blunt categories, as our previous chapters already have shown. Much of the rest of the book will be about understanding more fine-grained relationships between electoral systems and party systems. As for executive types, their ranges

[7] Examples include Lijphart (1994), Cox (1997), Mozaffar et al. (2003) – but see Filippov et al. (1999).

TABLE 4.2 *Mean values of effective number of parties (seats and votes) by electoral system and executive binary categories*

	Parliamentary	**Presidential**
Effective Number of Seat-winning parties (N_S)		
FPTP mean	2.14	1.96
IQR	1.74–2.35	1.92–1.99
Number of elections	159	37
PR mean	4.23	3.60
IQR	3.22–5.04	2.27–3.90
Number of elections	261	63
Effective Number of Vote-earning parties (N_V)		
FPTP mean	2.75	2.10
IQR	2.05–3.04	2.02–2.11
Number of elections	157	37
PR mean	4.77	4.32
IQR	3.59–5.67	2.56–4.78
Number of elections	261	63

IQR = Interquartile range (the range between the 25th and 75th percentiles)

of N_V and N_S overlap heavily. While it is true in all categories (seats or votes, FPTP or PR) that the mean values are lower in presidential systems, the differences are small.[8] We will see later in the book that there is not a systematic impact of presidentialism on the (effective) number of parties in the assembly, once the electoral system variables are specified correctly.

The effective number of parties based on votes (N_V) almost always exceeds the one based on seats (N_S), as we see from the averages in Table 4.2, although exceptions occur. That the N_V is usually larger than N_S, but not by much, proves to offer an essential building block for developing a logical model of votes distribution. We develop that model in Chapter 8.

PROPORTIONALITY: BASIC INDICES

Now we turn our attention to measuring proportionality – or more precisely, *deviation* from the ideal of perfect proportionality.[9] In practice, what this

[8] An important caveat here is that the sample of presidential systems with FPTP is almost entirely one country, the United States. (There are two elections in Ghana and two in Sierra Leone included.)

[9] We do not mean "ideal" in a normative sense, but a scientific one. It is a standard against which to measure, regardless of one's preference over whether full proportionality is "good" or "bad" as an outcome.

means is the deviation of seat shares from vote shares of parties.[10] Two ways to measure deviation from PR have dominated. Both start with the difference between votes and seats, for each party, but they differ in how they process these differences. Loosemore and Hanby (1971) introduced into political analysis the index of deviation that we'll designate as D_1, following the systematics of Taagepera (2007: 76–79). For deviation from PR, it is

$$D_1 = 1/2\sum|s_i - v_i|.$$

Here s_i is the i-th party's seat share, and v_i is its vote share. The index can range in principle from 0 to 1 (or 100 percent). Note that $|s_i - v_i| = |v_i - s_i|$ is never negative. D_1 dominated until Gallagher (1991) introduced what we'll designate as D_2:

$$D_2 = [1/2\sum(s_i - v_i)^2]^{1/2}.$$

It has often been designated as the "least squares" index, but this is a misnomer. The index does involve squaring a difference but no minimization procedure so as to find some "least" squares. D_2 can range from zero to one (100 percent), but whenever more than two parties have nonzero deviations the upper limit actually remains below one – an awkward feature (Taagepera 2007: 79–82). When only two parties have nonzero deviations, the one gaining what the other is losing, then D_1 and D_2 have the same value. But when more than two parties have nonzero deviations, then D_1 is bound to be larger than D_2. In sum,

$$D_1 \geq D_2.$$

It possible, though rare, that one of these indices increases while the other decreases, from one election to the next. For further detail on various indices of deviation from proportionality, see Taagepera (2007: 65–82) and Taagepera and Grofman (2003).

Gallagher's D_2 rapidly displaced D_1 during the 1990s as the more widely used index,[11] despite grounds for doubting whether it is the best of the various measures (Taagepera and Grofman, 2003; Taagepera 2007: 76–78). In this

[10] We can apply the same mathematical format as deviation from PR to other features of potential interest: volatility of votes from one election to the next; the extent of ticket splitting, when voters have more than one ballot; deviations of individual district magnitudes from the system mean (see Chapter 16). All these phenomena deal with measuring *deviation from a norm* (whether it be seat shares equal to vote shares, zero ticket splitting, or equal magnitudes).

[11] It even entered popular media discourse in Canada in late 2016, when a parliamentary committee charged with considering alternative electoral systems released its report, and the Minister of Democratic Institutions mocked the formula in making her case against the committee's recommendation for greater proportionality. See "The problem with Maryam Monsef's contempt for metrics," *McCleans's*, December 3, 2016, www.macleans.ca/politics/ottawa/the-problem-with-maryam-monsefs-contempt-for-metrics/ (accessed December 15, 2016).

book's Chapter 9 we introduce a logical model that accounts for deviation from PR, defined as D_2.

Empirical Patterns of Deviation from PR

It may seem obvious that features of the electoral system would affect deviation from proportionality. After all, some electoral systems are called "proportional representation" for a reason! We can do better than just say that proportional systems have lower deviation from proportionality than majoritarian systems, such as FPTP. **Figure 4.1** shows a scatter plot of those countries from our dataset that have "simple" electoral systems. They are plotted according to their mean district magnitude on the horizontal axis and deviation from proportionality,

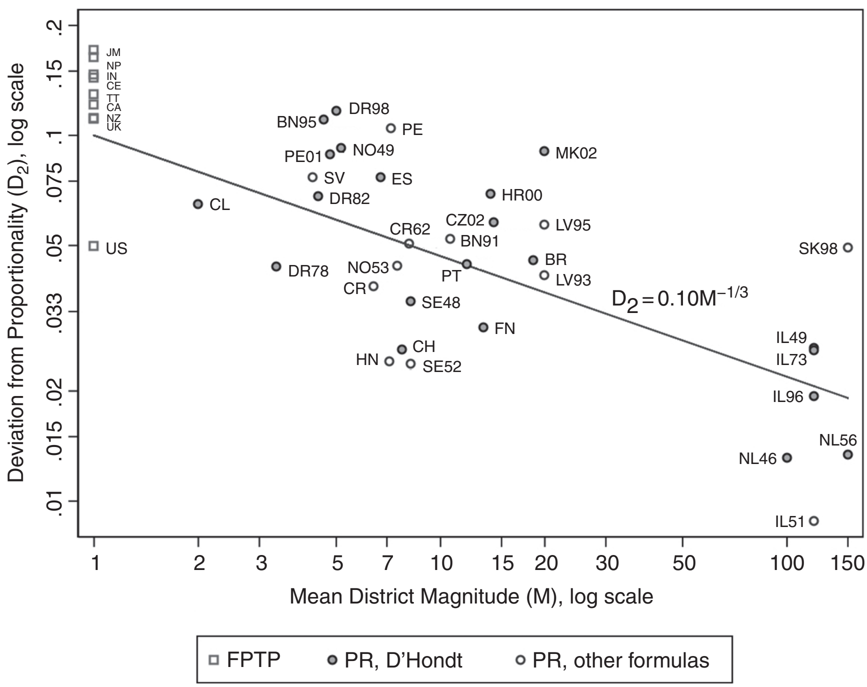

FIGURE 4.1 Mean district magnitude (M) and deviation from proportionality (D_2)
Country abbreviations: BN=Benin; BR=Brazil; CA=Canada; CE=Ceylon; CH=Switzerland; CL=Chile; CR=Costa Rica; CZ=Czechia; DR=Dominican Republic; ES=Spain; FI=Finland; HN=Honduras; HR=Croatia; IL=Israel; IN=India; JM=Jamaica; LV=Latvia; MK=Macedonia; NL=Netherlands; NO=Norway; NZ=New Zealand; PE=Peru; PT=Portugal; SE=Sweden; SK=Slovakia; SV=El Salvador; TT=Trinidad and Tobago; UK=United Kingdom; US=United States. If a two-digit number follows the country abbreviation, it is to indicate the first year of a period under a different electoral system from others depicted for the same country.

D_2 (Gallagher) on the vertical axis. Both axes are on logarithmic scales. Where a country abbreviation appears with a two-digit number immediately following, it is due to some change in its electoral system.[12] Figure 4.1 differentiates systems according to allocation formula: squares for FPTP (M=1) and circles for PR (mean M>1). In addition, if the PR formula is D'Hondt, the circle is filled in, but for other formulas, it is hollow.[13]

There is clearly a strong relationship, and it can be described approximately by the solid black line, which corresponds to:

$$D_2 = 0.10M^{-1/3}.$$

The relevance of the slope, –1/3, will be explained in Chapter 9.

Figure 4.1 and its symbol patterns allow us to see that the specific formula in a PR system makes surprisingly little overall difference to the relationship between D_2 and mean M.[14] It is evident that magnitude really is the "decisive factor," as we called it in Taagepera and Shugart (1989a).[15] It is also noteworthy how low D_2 is in the United States, compared to other FPTP systems, although its use of M=1 results in higher deviation from proportionality than about half the PR systems.

CONCLUSION

Two of the most important standards conventionally used to assess electoral systems and democratic competition are the number of parties and proportionality. It is common to speak of a "two-party system" or a "multi-party system," and even casual observers of politics have some idea in their minds about how democracy differs under these categories of party system. Moreover, proportionality is such a central indicator of outcomes that an entire class of electoral system is known by the term, proportional representation, to distinguish it from other, typically quite disproportional systems such as First-Past-The-Post.

[12] We follow the definition of a new "electoral system" that Lijphart (1994) established: a change in formula or a change of at least 20 percent in mean district magnitude, assembly size, or legal threshold.

[13] Most of the non-D'Hondt PR formulas shown are either Hare quota with largest remainders or Ste.-Laguë divisors.

[14] Nonetheless, see Carey (2018) for some examples of where formula has made a substantial difference in election outcomes in some newer democracies where the vote was highly fragmented after the plurality party.

[15] We see some impact of formula, independent of M: The lowest value of D_2 in Figure 4.1 is for Israel when it used Hare quota with largest remainders (1951–1969); Israel's D_2 values are strikingly higher in the periods when it has used D'Hondt. (The other separate indicators for Israel all refer to changes in legal threshold. For detail, see Chapter 5 and Hazan et al. (2018); the impact of thresholds will be discussed in Chapter 15.) In addition, the data points marked "SE48" and "SE52" show the impact of that Sweden's change from D'Hondt to Modified Ste.-Laguë.

In this chapter, we took these concepts and showed how political scientists develop measures that summarize the number of parties and proportionality for any given constellation of votes and seats. The measures themselves are not new. The index for the effective number of parties, as a summary measure of the fragmentation of party support in either votes or seats, has been around since 1979. Various others had been proposed earlier, and new ones have been proposed since. All share the basic aim of indicating, with just one number, how closely a party system approximates a "two-party system" or just how "multi-party" it is. No single index ever can sum up everything we might want to know about something like a party system, where there may be many parties of wildly unequal sizes. Yet the effective number of parties has stood the test of time. More importantly, as subsequent chapters will show, it lends itself well to logical modeling that allows us to understand how institutional features of the electoral system shape party competition. Similarly, various measures of deviation from proportionality have been around for decades by now. Here, too, there is no perfect measure, but Gallagher's index has become widely used now, and we will see that it, too, has some promise for the development of logical models.