CONSTRUCTING THE NUMBER OF PARTIES

Patrick Dunleavy and Françoise Boucek

ABSTRACT

Forming an idea of the number of parties competing at elections or
winning seats in legislatures is fundamental to disaggregated approaches
to mapping party systems. We set out a method for systematically
relating the behaviour of any 'number of parties' index to the size of the
largest party's vote and the numbers of parties in competition. This
approach shows that the 'effective number of parties' (N_2) can confuse
real changes in party competition with mathematical quirks in the way
that the index is calculated. We also demonstrate that N_2 (and its main
rival the Molinar index) behaves in hard to predict and anomalous ways
under some configurations of party support. We conclude that the
Molinar index should not be further used, and that the N_2 score's
behaviour can create problems in quantitative applications. Even in less
formal historical or comparative analyses N_2 always needs to be
carefully interpreted. There is no 'perfect' measure of the weighted
number of parties, but averaging N_2 scores with a simple measure of
largest party predominance ($1/V_1$) produces a highly correlated measure
(Nb), but one with lower maximum scores, less quirky patterning and
a readier interpretation. A more radical solution is to 'spatialize' N_2 (or
Nb) scores, which allows analyses to take more account of variations in
the party competition conditions lying behind any given index number.

KEY WORDS ■ objective number of parties ■ counting rules

'Normal science' processes work by the accretion of knowledge in cumula-
tive, coral-reef fashion, allowing a scientific consensus to emerge which can
sustain further work at the frontiers of knowledge, without constant foun-
dational critiques disturbing the core concepts and theories of the discipline.
In the study of political parties, the effective number of parties index has
gradually reached a high level of acceptance since its first exposition by
Laakso and Taagepera (1979). The index is a measure of the level of concen-
tration in political life which assigns more influence to large parties and

screens out very small parties in its computation: 'The assumption in the comparative politics literature has long been that some kind of weighting is necessary' (Lijphart, 1994: 67). Influential authors such as Lijphart (1984) advocated the general adoption of the measure, and 10 years later he described it as 'the purest measure of the number of parties' (Lijphart, 1994: 70). He also claimed (p. 68) that: 'In modern comparative politics a high degree of consensus has been reached on how exactly the number of parties should be measured.' Lijphart's confidence in the measure has continued to grow: 'The problem of how to count parties of different sizes is solved by using the effective number measure' (Lijphart, 1999: 69).

However, we show here that the effective number of parties is a somewhat flawed index, whose use in quantitative analysis can create problems. And as it is commonly interpreted in most of the electoral and party analysis literature, the measure (and the wider family of indices of which it is a part) can be misleading. The root of these problems is primarily that the political scientists who devised or advocated the index never gave a systematic *experimental* account of how its results were patterned across the full range of possible empirical outcomes. We give here a comprehensive exposition of how the effective number index behaves with changes in two key variables: (i) the level of support for the largest party (V_1); and (ii) the number of observable parties in competition.[1]

1. How the Effective Number of Parties Index Works

Early authors taking a dimensionalized approach to party systems analysis saw their task as easily accomplished by quickly constructing artificial numerical measures, devoid of any ready intuitive meaning (see, for example, Lane and Ersson, 1987: 154–79; Mayer, 1980a, b; Wildgen, 1971). But these measures do not relate in any obvious ways to 'ordinary language' descriptions. They can seem arbitrary and author-specific. No professional consensus developed behind their usefulness and they were little cited. So, with a few exceptions, such as the party 'fractionalization' index devised by Rae (1967: Ch. 3), the 'non-intuitive' indices did not widely influence political science thinking about party systems. The great achievement of the effective number of parties measure was to offer a simple to calculate and more easily understandable index of the number of parties in competition which is neither dependent on just the largest party's vote ($1/V_1$) nor distorted by alterations in the numbers or vote shares of very small parties (Laakso and Taagepera, 1979; Taagepera and Shugart, 1989: Ch. 8). The index accords some weight to all parties. But it weights the largest parties most, while small parties (with support under around 10 percent) count for very little and tiny parties (under 1 percent) count hardly at all. The measure can also be simply described as 1 divided by the sum of the squared decimal shares of the vote for (or seats won by) each electoral

party. Indeed, that is how the index is generally presented (Lijphart, 1994: 68; Taagepera, 1999):

$$N_v = \frac{1}{\sum\limits_{1}^{x} v_i^2} \qquad (1)$$

where vi denotes the share of the votes going to each party i, and the notation above and below the summation sign shows that it covers all parties from the largest (V_1) to the smallest (Vx).

Handy as this simplified representation may be, it disguises the fact that the normal effective number is in fact part of a more complex family of indices, whose generic formula is:

$$N_a = \left[\sum\limits_{1}^{x} v_i^a \right]^{1/(1-a)} \qquad (2)$$

Rephrasing this in operational terms (for example, the procedures which need to be followed by someone using a spreadsheet or SPSS to compute empirical scores) all these indices have the following steps:

- take the decimal vote shares for all of the parties and raise each to power a;
- add up all the raised numbers; then
- raise the resulting summed number to 1 divided by $(1 - a)$.

The conventional effective number index is actually the special case N_2 where $a = 2$: in this case the power term $1/(1 - 2)$ reduces to -1, meaning that we divide 1 by the squared vote shares of all the parties. All of the family are mathematically constrained so that: 'If all components have equal shares, then the effective number must be the same as the actual number' (Laakso and Taagepera, 1979: 5), which the authors regard as the first requirement for any 'rational' index of this type. By 'actual number' here they mean the observable number of parties, counting each party as 1 irrespective of its size. This feature has also been a key peg used by later exponents to help explain the index; for instance, Lijphart (1999: 68) notes that: 'In all cases where all the parties are exactly equal, the effective number will be the same as the raw numerical count.' This important requirement creates whole-number 'anchor points' at 2, 3, 4, 5 parties and so on, to which all the indices must converge periodically when there is minimum fragmentation of parties.

Exponents of N_2 argued that other indices in the effective number family could also be useful measures (Laakso and Taagepera, 1979; Taagepera and Shugart, 1989). But there is no evidence of their being picked up in the literature. Some indeed are almost pointless, and others cannot be calculated in any practicable or intuitively understandable way using equation (2) above, which seriously impairs their value in empirical research.

Taagepera (1999: 408–9) suggests that N_0, N_1 and N_4 are all useful. But N_0 is just the count of observable parties, and no information is added by working through equation (2) above with $a = 0$. For N_1, setting a to 1 itself means that the power term $(1/1-1)$ becomes zero, giving an overall answer of zero whatever the party vote shares. So here the index must be calculated instead as $N_{\rightarrow 1}$ (for example, making $a = 0.99$), which of course has no ready intuitive interpretation. And we show below in Figure 1 that the maximum fragmentation line for $N_{\rightarrow 1}$ seems to accord poorly with our intuitive estimations, counting up to 3 parties when the largest party has 80
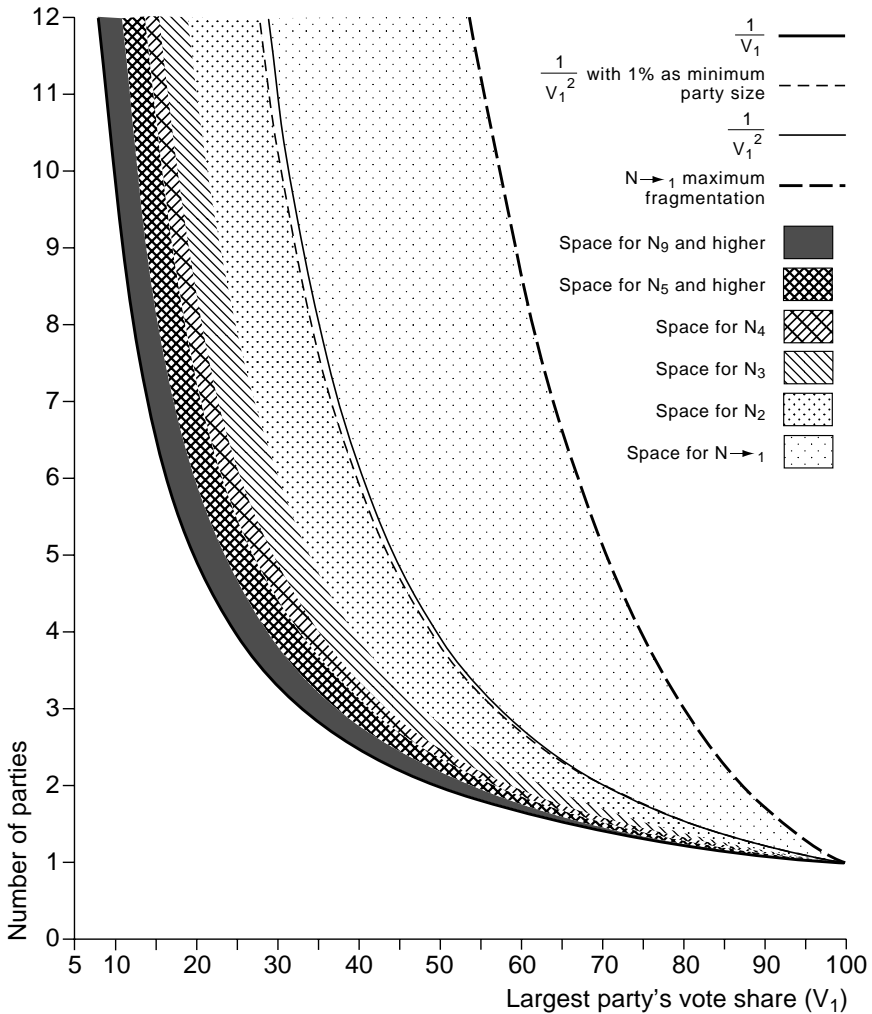

Figure 1. The minimum and maximum limits of the space for the 'effective number of parties' families of indices

percent support, and more than 12 when V_1 has majority support. N_∞ does have a ready mathematical interpretation, since in practice it reduces to 1 divided by the (unsquared) vote share of the largest party (that is, $1/V_1$). But note that for virtually all normally equipped empirical political scientists N_∞ is completely incapable of calculation using equation (2). Because conventional computers or PCs cannot compute forever, raising every party's decimal vote share to power infinity (or indeed to any power above around 800) yields a result indistinguishable from zero: raising the sum of multiple zeros to power infinity then still yields zero. So why employ the non-operationalizable concept N_∞ when the simple to understand and compute measure $1/V_1$ does the job so much better?

Figure 1 shows the basic confines of the space available with different members of the effective number of parties family. The $1/V_1$ curve marks the lower boundary, when there is no space at all for fragmentation of the opposition to have any effect on the score. Of course, it intersects the whole-number 'anchor point' scores recorded under all the effective number indices when parties split the vote exactly equally between them. The maximum fragmentation lines for any value of a are identified by assuming an infinite number of infinitely small opposition parties (that is, V_2 to V_x). At any value for power a above 1 their tiny initial vote shares are consequently raised into negligibility, giving upper boundaries where V_1^a is raised to $1/(1 - a)$. The very high upper limit for $N_{\rightarrow 1}$ is clear, and the N_2 upper bound (at $1/V_1^2$) is also much higher than that for any other effective numbers index. The outer limits for N_2, N_3, N_4, N_5 and so on progressively retreat back towards the $1/V_1$ line, although in rapidly decreasing jumps. When a reaches 9 the figure shows that there is already very little space left for opposition fragmentation to influence the results. Of all the possible effective number indices, only setting a to equal 3 or 4 (or just possibly 5) seems to offer any advantage compared with N_2. The maximum lines for these variants are not as far out as for N_2 (at $1/V_1^2$). But at the same time these spaces are not so close to the $1/V_1$ line that changes in the size and composition of the opposition parties have little or no influence upon the effective number score.

At first sight the boundary curves in Figure 1 look well behaved as support for the largest party varies. But we next need to consider how the same scores are configured with different numbers of parties in competition. It may seem odd that such an established and widely used index as N_2 should still need additional clarification. But in describing the measure, Taagepera and other authors primarily concentrated on its algebraic form, how the index works in some erratically selected hypothetical vote share situations, and how it scores a small number of empirical vote and seat shares in different countries. For instance, in his most recent discussion, Taagepera (1999) offers just 12 illustrative vote combinations out of the many possible to show how his family of indices operates. But what we really need is to understand empirically how an index performs across the full range of all

possible situations – what the range of *potential* variation is for N_2 – rather than just how the index works in isolated instances.

To achieve such a picture concisely, we discard the notion of an infinitely large number of parties, and focus instead on situations where all parties are a certain 'relevant' size, namely 1 percent of votes or seats or more. This simplifying step removes some of the difficulties of counting all observable parties, by assuming away very small units. And it makes only a small difference to the results in most cases: for instance, in Figure 1 the dashed grey line tracking just inside the theoretical $1/V_1^2$ curve shows the same curve with the 1 percent limit applied. Assuming only relevant parties also has a strong rationale when computing N_2 or other indices for seats data, since all legislatures or committees have minimum size thresholds (set by 1 over the number of seats), and many have much higher legal minimum thresholds for parties to win representation.

We next define scenarios for the minimum and maximum fragmentation of opposition parties (V_2 to V_x) which set logical lower and upper limits to the values that can be registered under any number of parties measure, at any given value of V_1:

- (*The minimum number of relevant parties* occurs where so far as possible the second and subsequent parties (bar one) obtain the same share of the vote as V_1 minus the smallest possible amount. So where x is the number of empirically observable parties, $V_1 \oplus V_2 \oplus V_{x-1}$ (where V_{x-1} is the second smallest party), and then Vx mops up any remaining vote share. (For example: with V_1 at 61 percent, V_2 is 39 percent. With V_1 at 48 percent, V_2 is at the same level less a small decrement, and V_3 is 4 percent. With V_1 at 23 percent, V_2, V_3 and V_4 are also at this level, and V_5 is at 8 percent.)
- (*The maximum number of relevant parties* occurs where there is the greatest possible gap between V_1 and all the opposition parties on the minimum size of 1 percent. Thus if V_1 is 61 percent, there are 39 other parties each on 1 percent. Far fewer empirical observations may lie close to the maximum fragmentation line than the minimum fragmentation line, because the high coordination costs for voters and political leaders in operating with so many parties produce strong pressures for party coalescence (Cox, 1997).

Figure 2 shows how the effective number of party score (N_2) varies across different levels of V_1 as the number of relevant parties increases from 2 to 3, 4 and so on up to 8 parties. (Beyond that the spaces continue the same pattern in a progressively more bunched together way, which we do not show here because it would make the diagram too complex.) The $1/V_1^2$ line shows the maximum fragmentation line for a very large number of parties (with a minimum party size of 1 percent). Laakso and Taagepera's treatment of equal components means that substantial parts of the potential space of scores under N_2 shown in Figure 1 cannot actually be accessed.
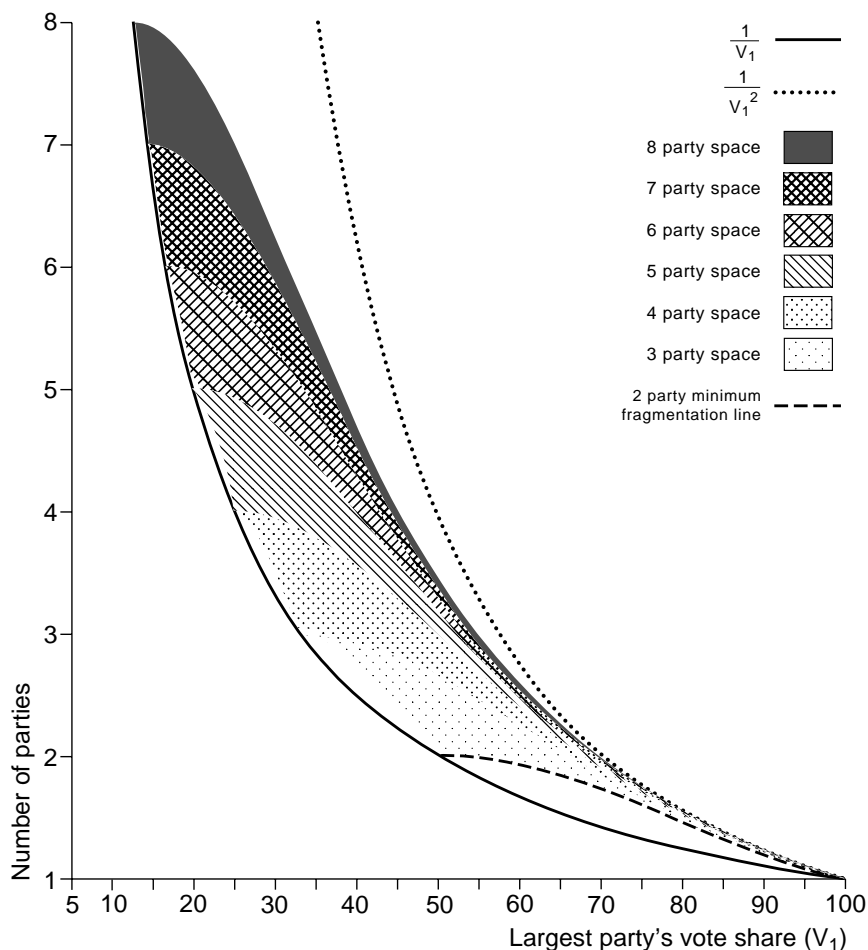
Figure 2. How the shape of the space for the effective number of parties varies with the number of relevant parties

And the realizable space of scores actually builds up in a fairly complex 'bat-wing' pattern as the number of relevant parties increases. With only two parties the space of scores is just a single, upward-bowed line (since party vote shares cannot differ between minimum and maximum fragmentation here). The line (shown dashed towards the bottom of Figure 2) starts at 1 when the largest party has 100 percent support. It then rises well above the $1/V_1$ line as V_1 declines to around 65 percent, from where it curves back to join the $1/V_1$ line at the first 'anchor point' where both parties have 50 percent each.

With three relevant parties the available space changes from a line into an area, shown in the lightest shading in Figure 2. The bottom of this area

is defined by the minimum fragmentation line with three observable parties. It runs just above the two-party line, while V_1 has majority support (and hence it is not shown separately in Figure 2). When V_1 reaches 50 percent, this minimum line does not quite touch the $1/V_1$ line here, but instead loops sharply up again to curve back and join the $1/V_1$ line only at the second anchor point, when three parties garner exactly a third of the vote each. The upper boundary of the three-party area (the top edge of the lightest shaded area in Figure 2) is given by the maximum fragmentation line with three relevant parties. It initially rises to reach its maximum height above the $1/V_1$ line when V_1 is at 50; thereafter it curves upward much more shallowly so that it also converges on the three-party anchor point (where all parties have equal vote share).

Figure 2 shows that the pattern set by the three-party space repeats as the number of parties in competition increases. Each successive increase in the number of relevant parties adds another large segment to the bat-wing and includes almost all of the space for the previous number(s) of parties. So, the space of scores with four observable parties runs almost from the bottom margins of the three-party space (excepting only the three-party minimum fragmentation line itself) to the top of the four-party space (defined by the four-party maximum fragmentation line). Similarly, the space of scores for a situation with 8 observable parties covers all of the shaded areas for 7, 6, 5, 4 and 3 parties shown in Figure 2, except for a succession of slithers off the bottom and left-hand margins of the three-party space, as the minimum fragmentation line for higher numbers of parties creeps up in small increments (which are too small to represent here). Note that tiny changes in the number of relevant parties in competition can exercise an enormous influence on the space of the N_2 index available. The addition of an extra party (attracting just 1 percent support in this analysis) always radically transforms the shape of the region within which an N_2 score can be located.

To see how the N_2 index operates empirically with the support levels for the largest parties that are common in established liberal democracies (between 20 and 60 percent), Figure 3 shows how 102 post-war election results from 7 countries are distributed across the 'bat-wing' areas. The countries are the United Kingdom and Canada using plurality systems; Ireland, Sweden and Italy using list PR or STV systems; and Japan and France using intermediate systems. The space for four or fewer relevant parties is populated almost entirely by the UK, Canada and Irish results, although a handful of scores for Japan also squeeze in here. The results for the other four countries using PR systems are concentrated in the spaces added on when there are five, six or seven relevant parties.

A number of possible problems with N_2 are evident here. First, the lower boundaries of the spaces for small numbers of parties (that is, their minimum fragmentation lines) run well above the $1/V_1$ curve when the largest party has majority support. So it is impossible for any electoral
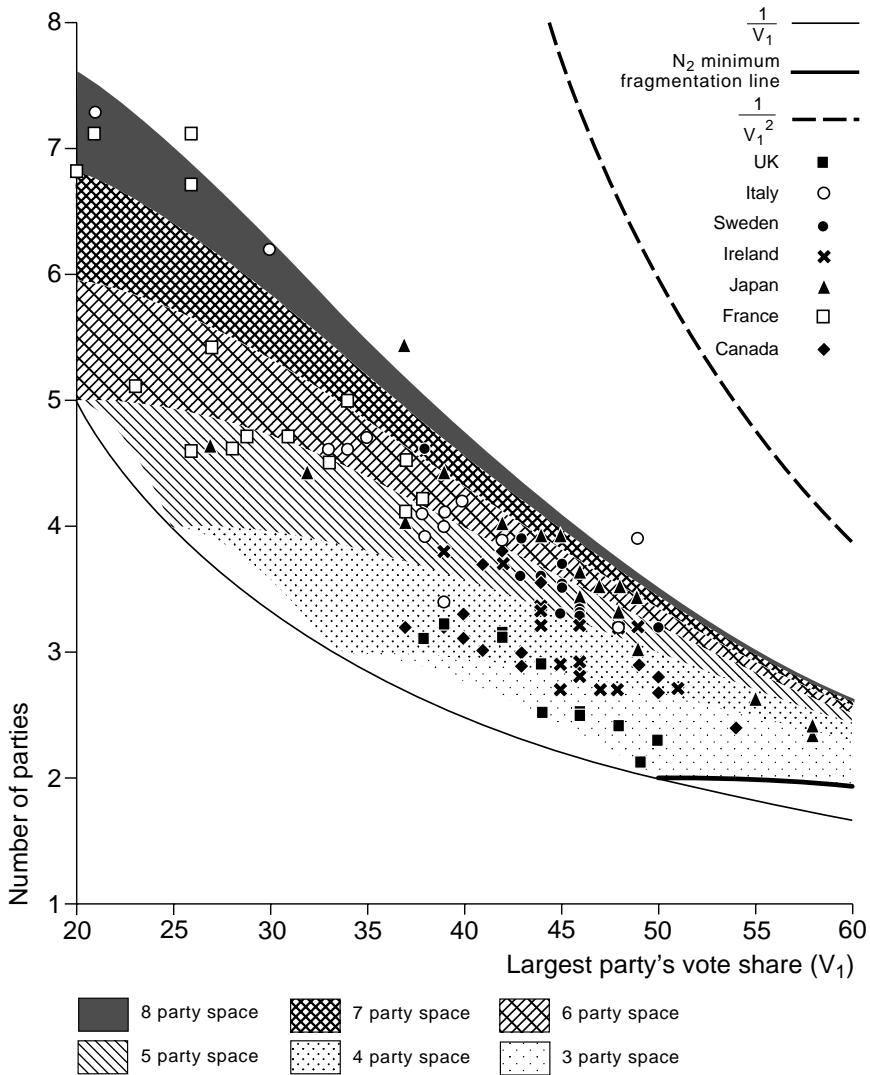
Figure 3. How the post-war election results for seven liberal democracies are distributed across the space of the effective number of parties index

system to get an N_2 score of *less* than 2 unless the largest party's vote share V_1 is very high, nearly 60 percent. (In fact, when V_1 is at 60 percent the lowest possible N_2 score is *above* the highest possible N_4 score.) Second, because N_2 accords so much more emphasis to opposition fragmentation than any other index (except N_1) this effect is further compounded. It is difficult to generate N_2 scores of less than 2 even at much larger V_1 levels, so that Molinar (1991) argued that N_2 systematically overstates party

fragmentation, compared with our intuitive judgements. For instance, with V_1 at 70 percent and three opposition parties on 10 percent each, the N_2 score is still over 1.9 – even though Molinar regarded the largest party in this situation as clearly a dominant one, facing no viable competitors. Taagepera (1999: 500–1) seems to recognize that N_2 has a structural problem here. His 'largest component approach' concedes that empirical analysts should not just consider N_2 alone, but also look at the N_4 score for each situation (which we have seen reduces to the $1/V_1$ line). In practice, however, analysts seem to be enjoined to notice only that a $1/V_1$ score which is over 2 indicates majority support for the largest party, a fact which is just as visible from looking at V_1 alone.

Third, two of the N_2 scores for plurality rule countries' elections in Figure 3 are right on the lower boundaries of their party spaces (their minimum fragmentation lines), demonstrating that their numerical values can be conditioned just by the 'kink' effect here. As V_1 levels increase from left to right in Figure 3, these lower boundary lines drop very steeply as each of the anchor points is approached, but then fall hardly at all once they have been passed. For example, when V_1 is close to 50 percent the numerical scores under minimum fragmentation are:

V_1	44	46	48	49	**50**	51	52	54	56	58	60
N_2 score	2.49	2.33	2.16	2.08	**2.0**	1.999	1.996	1.99	1.97	1.95	1.92

The N_2 score falls by a hundred times as much when V_1 moves from 49 to 50 percent, as it does if V_1 moves from 50 to 51 percent. And it falls by 33 times as much when V_1 moves from 46 to 50 percent as it does if V_1 moves from 50 to 54 percent. Data accumulated close to any of the anchor points (with a varying number of relevant parties) will thus be contaminated by purely artefactual effects. Misleading results could be obtained especially if such data are correlated or regressed with other variables in a quantitative analysis. Note that such problems of artefactual variations in scores will almost certainly arise more often when N_2 scores are used to estimate the number of legislative parties or electoral parties in plurality rule systems (see Colomer, 2001: 70–116; Lijphart, 1999: 67–9, 77–8). The reductive effect of leader's bias and other effects under plurality rule squeeze the support for and representation of third and fourth parties, and hence will typically push these scores closer to the N_2 theoretical minimum for that number of relevant parties.

We noted above that switching to N_3, N_4 or another member of the effective number family with a low value of a could solve N_2's problem of allowing exaggerated maximum fragmentation scores. But this step cannot cure the problem of kinks around the anchor points. Figure 4 shows how the minimum fragmentation lines for all the low value effective number indices behave when V_1 has close to 50 percent support. For $N \rightarrow_1$ the change
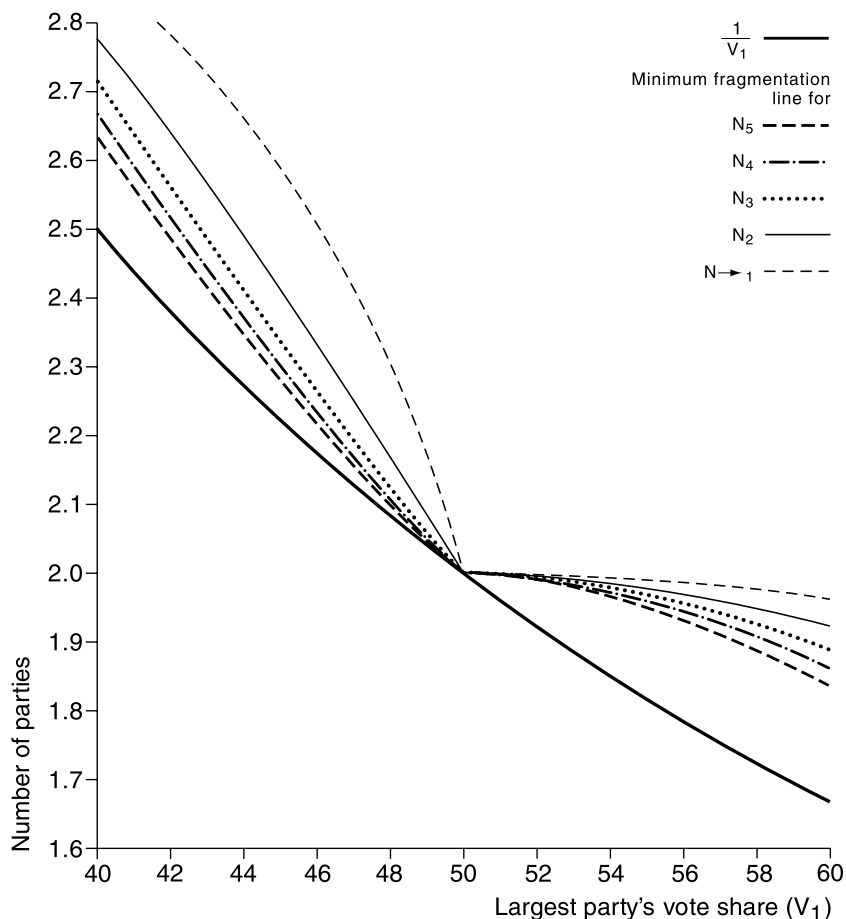
Figure 4. The behaviour of effective number of parties indices around the 50 percent anchor point, under minimum fragmentation conditions

of angle is even worse than for N_2, while effective number indices with values of a such as 3, 4 or 5 show little improvement on N_2. A smoother curve only appears with high values of a. But once a exceeds 9 or 10 the indices suffer because changes in the fragmentation of opposition parties make almost no difference to the number of parties score.

The effective number approach additionally confronts a generic problem of all such weighted indices, that any given N_2 score may be produced under widely differing conditions in terms of the numbers of observable parties in competition and the size of V_1. Comparing the raw numerical N_2 scores produced with different numbers of parties in competition and different levels of largest party support is potentially misleading. For instance, there are seven scores of 3.9 shown in Figure 3, but they occur across completely

different segments of the bat-wing areas. One is in the space added for five relevant parties, three in the space for six parties, another two in the space for seven parties, and one (for Italy) in the area (not shown in Figure 3) for 11 relevant parties. Simply treating all these scores as numerically equivalent ignores the fact that they are produced under very different conditions. These problems are most severe with scores which occur most commonly under 'normal' political conditions in liberal democracies: Figure 3 shows a concentration of V_1 levels between 35 and 50 percent and of N_2 scores between 2.5 and 4 parties.

3. Can We Improve on the Effective Number Score?

Despite these difficulties, the effective number approach has few rivals in constructing the number of parties. One important alternative, devised by Molinar, has been discussed only by a few specialists in the field and not widely taken up. In the Appendix below we show that despite the vigour of Molinar's critique of N_2 his index has many more deeply problematic features. We conclude from this analysis that the M index should not be further employed in political science. Does this then leave the field open for N_2 (or another close member of the effective number family, such as N_3)? We do not believe that it should because of the index-specific problems mentioned above – N_2's over-rating of the number of parties, especially when V_1 has majority support; its over-response to fragmentation of the opposition vote; and the probability of artefactual results introduced by the kinks around anchor points.

One possible way of coping with the limitations of the effective number of parties is to combine it with another index that can correct for its failing. Taagepera (1999: 503) remarks dismissively of this idea:

> One may harbour the illusion that by judicious combination of N_2 and N_4 (plus possibly something else) one might achieve a super-index that satisfies all desiderata. This is about as wishful as hoping to combine the mean and the standard deviation of a distribution into a single measure. Two numbers are inherently able to transmit more information than a single one.

This is not a very developed line of argument – after all, two inherently limited numbers are not much use to anyone, and if they can be combined more productively then they should be. Taagepera's stance seems to be motivated chiefly by an understandable desire to defend the N_2 index from criticism, and he actually uses the $1/V_1$ element only as a 'supplementary index [which] may explain some apparent anomalies or at least make us cautious' (Taagepera, 1999: 502). His position rests on the notion that N_2 is a defensible measure for use in quantitative analysis, and his lonely faith in the value of other effective number indices, for which there has been little or no take-up in the existing literature. By contrast we believe that the wider

effective number family has little to offer, and that continuing to use un-modified N_2 in particular in quantitative applications cannot be defended because of the defects set out here.

In our view, averaging N_2 scores with the $1/V_1$ score creates a simple but useful variant of the effective number index, N_b:

$$N_b = \left(\frac{\frac{1}{x} + \frac{1}{v_1}}{\sum_1 v_i^2} \right) \times \frac{1}{2} \tag{3}$$

The data demands of equation (3) are no greater than for the N_2 index, and N_b and N_2 are highly correlated with each other. Yet this straightforward modification has useful effects. Figure 5 shows the spaces for N_b with between 2 and 8 parties, with the top of each space defined by a maximum fragmentation line. It also includes the $1/V_1$ line and the overall maximum fragmentation line for N_b, given by adding $1/V_1$ and $1/V_1^2$ and then dividing by 2. (Again we assume a 1 percent floor for party sizes, as before.) The averaging of N_2 and $1/V_1$ creates much less curved lower bounds of the spaces (minimum fragmentation lines). And although there are still tran-sitions in their slopes around the anchor points, they are much less sharp than with N_2. The top boundaries of the spaces for different relevant numbers of parties (maximum fragmentation lines) are also considerably straightened out under N_b, without strongly visible curves close to their terminal anchor points. The overall maximum fragmentation line for N_b is appreciably lower than the $1/V_1^2$ line under N_2. In fact the N_b maximum fragmentation line runs quite close to but slightly above the N_3 maximum line shown in Figure 1. For instance, with V_1 at 60 percent, the maximum N_b score is more than half a party less than with N_2; and at 50 percent support the N_b upper limit is 3 parties, instead of 4 for N_2. Thus the N_b index delivers many of the same benefits in terms of more realistically denominated scores as N_3, but it avoids N_3's severe kinks around anchor points (which is evident in Figure 4).

Table 1 shows how the N_2, N_b and Molinar measures behave empirically across the 102 post-war election results graphed in Figure 3. The Molinar index (M) rates Japan and Sweden as the most concentrated party systems (reflecting its sensitivity to the gap between V_1 and V_2). In addition, M squashes the median number of legislative parties below 2 for both coun-tries (as well as for the UK, Canada and Ireland). In sharp contrast, N_2 portrays Canada, Ireland, Sweden and Japan as all clearly multiparty systems at an electoral level, with median scores over 3. This view seems hard to square with the widespread view that Sweden and Japan were dominant party systems for much of this period, or that Ireland and Canada had substantial periods of one-party predominance. By contrast, the N_b scores for every election fall in between the N_2 and M scores, so the index seems to avoid problems associated with generating extreme scores. Like
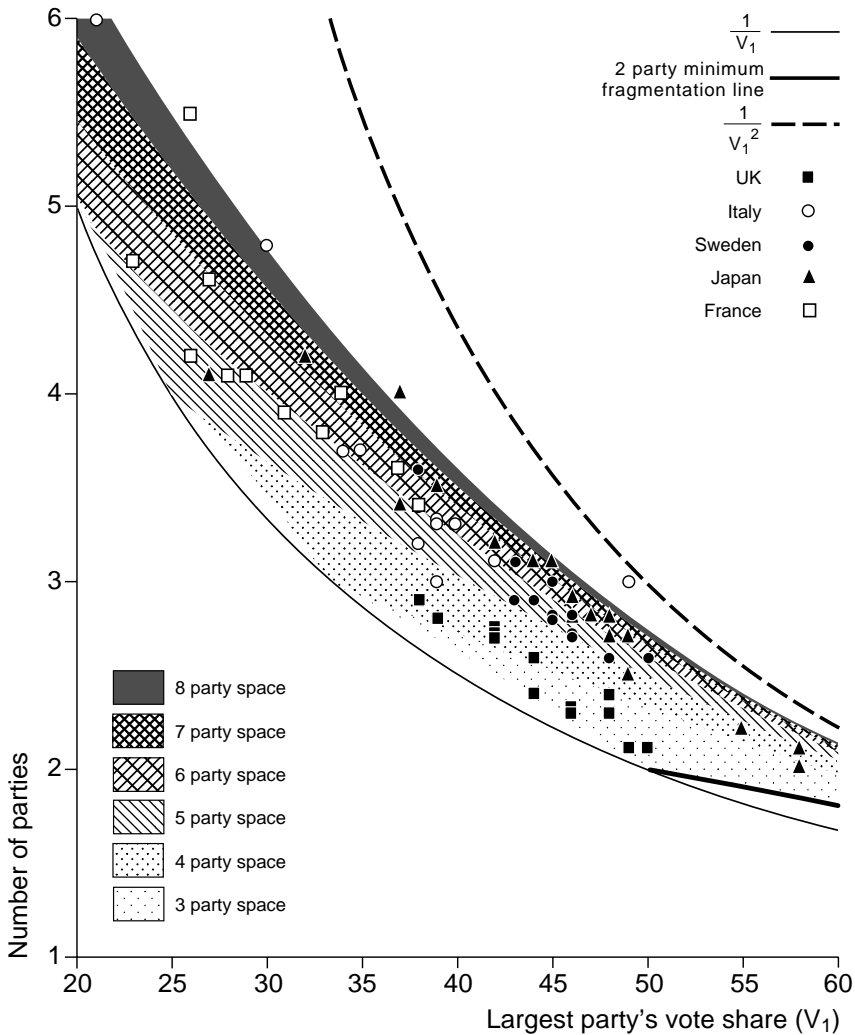
Figure 5. How the shape of the modified effective number of parties index (Nb) varies with the number of relevant parties

N_2, the N_b index yields scores above 3 for all the French and Italian elections. But in the other five countries N_2 gives scores of 3 electoral parties or more on 54 (out of 70) occasions, whereas N_b does so only 21 times. Similarly, the M index generates scores for legislative parties of under 2 in 49 of the elections considered in Table 2, compared to just 13 times for N_b.

Although the Molinar values in this table are always less than those under N_2, when the two sets of index values are used to calculate the 'relative reduction in parties' (RRP) measure, some surprising results emerge. Taagepera and Shugart (1989: 273) define RRP as:

Table 1. Comparing the effective number of parties (N_2), the N_b variant and the Molinar index scores across 102 post-war elections in 7 countries, 1945 to 1994

	Median scores for Votes			Median scores for Seats		
	N_2	N_b	M	N_2	N_b	M
UK	2.6	2.4	2.25	2.1	2.0	1.85
Ireland	3.2	2.6	2.1	2.8	2.4	1.9
Canada	3.0	2.7	2.3	2.4	2.15	1.75
Japan	3.5	2.8	1.9	2.9	2.3	1.6
Sweden	3.45	2.8	2.0	3.3	2.7	1.9
Italy	4.1	3.3	2.6	3.6	3.0	2.4
France	4.7	4.1	3.7	4.1	3.65	3.1

The table is ordered by the N_b index score for votes.

$$\frac{\left(N_v - N_s\right) \times 100}{N_v}$$

Because the Molinar index drastically squashes down legislative party scores towards 1, calculating RRP using M scores can produce *higher* RRP results than those generated by N_2 scores – a result which emerged on 56 out of 102 occasions in our data set, with differences which are often appreciable. For instance, in the 1988 Swedish election both N_2 and N_b record more legislative parties than electoral parties, and hence generate negative RRP scores of –8 and –4, respectively. Their message is that far from suppressing the range of choices expressed by the electorate, the Swedish system of legislative representation slightly expanded them. But M identifies only 1.7 legislative parties after this election, giving an RRP score of +15 (post-war Sweden's second highest RRP score). Similarly, for the 1981 French election,

Table 2. Distribution of 'spatialized N_2 scores' for 102 post-war elections in 7 liberal democracies

	Number of results in the lowest relevant party space – 1							Re-averaged 'spatialized' N_2 score	Median post-war N_2 score
	2 parties	3 parties	4 parties	5 parties	6 parties	7 parties	8+ parties		
UK	8	6	–	–	–	–	–	2.4	2.6
Canada	1	13	2	–	–	–	–	3.1	3.0
Ireland	–	3	7	6	1	–	–	4.3	3.2
Sweden	–	–	12	3	1	–	–	4.3	3.45
Japan	1	3	5	8	1	–	1	4.5	3.0
France	–	–	5	5	2	2	1	4.8	4.7
Italy	–	1	2	7	–	2	1	5.6	4.1

In the 8+ column the French and the Japan cases were for 8 relevant parties, and the Italian case for 10.

when the Parti Socialiste and Communists finally defeated the Gaullists and their allies after 23 years of unbroken right-wing rule, the seats score for N_2 is 2.7 and for N_b 2.2. But the M score is 1.4 and the Molinar RRP measure is +50 (post-war France's highest score). These cases strengthen our view that M is not a useful measure, but that the N_b variant of the effective number of parties has stable properties.

While the shift to N_b scores can cope with N_2's unrealistically high numbers, over-sensitivity to opposition fragmentation and many of the problems with kinks around anchor points, it does nothing to address the problem that the same N_2 scores can be produced under very different party competition conditions. One way of dealing with this last point would be a method for setting empirically observed N_2 scores against the background of the bat-wing shapes shown in Figure 3 and then summarizing the results. We suggest that for any given N_2 (or N_b) score and V_1 combination it is useful just to record the lowest party space that it could fall into (irrespective of the number of relevant parties actually involved in the situation) and subtract 1. We call these simplified and transformed results 'spatialized' N_2 (or N_b) scores. To see how they work, consider the seven N_2 scores of 3.9 included in Figure 3 and shown in more detail below:

N_2 score	3.9						
V_1	38.3	41.9	42.4	43.2	43.9	44.6	48.5
Spatialized N_2 score	4	5	5	5	6	6	10
Country and date	Italy	Ireland	Italy	Sweden	Japan	Japan	Italy
	1979	1948	1958	1988	1949	1976	1948

We noted above that these results occurred in very different situations – in fact there is more than a 10 points difference between the lowest and highest V_1 levels underpinning these identical scores. To spatialize the scores, we check from Figure 3 the lowest number of parties in competition that could have produced a 3.9 with that level of V_1 and then subtract 1, giving the results shown in the third row. For instance, the two 3.9's produced by V_1 levels around 42 percent fall within the 6 party space in Figure 3, and are consequently coded $(6 - 1) = 5$. The 7 spatialized versions of $N_2 = 3.9$ here range between 4 and 10, graphically illustrating the non-comparability of these scores across different situations.

Table 2 shows the spatialized N_2 scores for the same 102 elections across 7 countries shown in Figure 3. Their distribution seems to accord well with our intuitive views of different party systems. Taking a re-average of spatialized N_2 scores for each country gives a new kind of average number for the underlying number of parties in that party system. Table 2 demonstrates more clearly than looking at the raw N_2 scores themselves the differences between the plurality rule systems in the UK or Canada and the list PR systems of Sweden or Italy. And it seems to characterize well the different

electoral systems of post-war Japan and France under the fourth and fifth republics. For three countries out of seven (the UK, Canada and France), the spatialized score average is virtually the same as the median N_2 score over this period. But in the other four countries here, the average spatialized N_2 scores are around 1 to 1.5 parties higher than the median N_2 scores, suggesting that they differentiate again between countries which the normal N_2 data represent as similar because of the index's over-response to fragmentation of the opposition. Note that the spatialization procedure does not do anything to control N_2's over-rating of the number of parties. It increases most high-scoring countries' numbers, so that their differences from low-scoring countries emerge more clearly, but it does not reduce N_2 scores overall. Substituting N_b scores as the base for the spatializing operation controls can yield useful insights while using lower base numbers. It is important to stress that for N_b scores as much as the standard effective number measure, any given score must be carefully interpreted against the background of the total number of parties in competition, the largest party's level of support and the level of fragmentation in the opposition vote (using the areas shown in Figure 5).

Conclusions

For a dimensionalized analysis of party systems to succeed, we must be able to construct measures of the number of parties across polities in a robust quantitative way, avoiding odd mathematical effects, and producing scores which fit reasonably well with our 'ordinary knowledge' ways of describing party systems (Lindblom and Cohen, 1979). The effective number of parties measure does not deliver a reliable or consistent relationship between changes in the largest party's share of the vote and the index number shown. The N_2 index behaves jerkily when party fragmentation is low, and yet accords more influence to fragmentation of the opposition than any other member of its family (except $N^{\rightarrow}_1$). It will often produce implausibly high scores for party systems where the largest party is in a strong governing position. Entering N_2 as an independent variable in quantitative analyses could produce artefactual effects or disguise real ones, especially when looking at legislative parties or at voting data for plurality systems. We also show in Annex I that the Molinar index has major quirks. It reverses the 'direction' of judgements in the N_2 index at high levels of V_1, recording highly fragmented party systems as having fewer effective parties: but it then ceases to behave in this way at lower levels of V_1. The Molinar index also systematically understates the numbers of parties in situations where one party has majority support (a problem especially in measuring legislative parties). By contrast to both these earlier indices, the N_b variant of the effective number of parties is a composite measure that yields more stable and readily interpretable results.

There is also a more general lesson for comparing across electoral systems in handling both the N_2 and N_b indices. The spaces within which index numbers can vary against V_1 change radically with increases in the number of relevant parties. So the meaning of any N_2 or N_b number can only be usefully determined by considering where it is located in the space of possible scores, given the level of support for V_1 in that particular situation. Even for analysts undertaking work on a single country or election sequence, we would recommend plotting the empirical scores obtained under N_2 or N_b against the basic spaces of areas for different numbers of relevant parties shown in Figures 2 or 5. This step can aid 'intuitive' thinking about the patterns shown and is a sensible precaution before taking on trust numbers which may reflect quirky index features. For more comparative, quantitatively based analyses the risk of fetishizing N_2 or N_b numbers without really appreciating their meaning is greater. We have suggested a method for 'spatializing' N_2 or N_b scores which can provide an additional stream of data which recognize where these score numbers are located against the fundamental spaces of their index.

As the movement continues away from older typologies of party systems and towards a more empirically sensitive description of party systems, correctly constructing the number of parties remains a very important issue (Ware, 1995). Dimensionalizing party systems with multiple indicators remains a promising agenda for research, but we should proceed more sceptically than in the past. There is no perfect number of parties index, but we have set out reasons why N_b scores are preferable to raw N_2 numbers in our view, and why spatialized N_b or N_2 scores are better still. Finally, we have used a basic method here of looking at the spaces within which index scores are feasible, defined by minimum and maximum party fragmentation conditions at varying levels of V_1 and under different numbers of relevant parties. We believe that this approach offers a valuable way of comprehensively evaluating the properties of any new party concentration measures (and indeed other indices) which may be proposed in future. It is vital that the behaviour of indices under the full range of possible conditions is systematically mapped from the outset, instead of years after they were first proposed.

Appendix: The Molinar Index

Molinar proposed an alternative formula to the effective numbers measure, calculated by multiplying the N_2 score by an additional term and then adding 1. This additional term is the sum of the opposition parties' squared vote shares expressed as a proportion of the sum of all the parties' squared vote shares. Formally:

$$M = 1 + \left(\frac{1}{\frac{1}{x} \sum_{1}^{x} v_i^2} \times \frac{\sum_{1}^{x} v_i^2 - V_i^2}{\sum_{1}^{x} v_i^2} \right) \qquad (c)$$

To show how his index performed, Molinar's original paper computed scores for 389 hypothetical vote combinations with different levels of V_1, and he also included small amounts of country data. He demonstrated that M scores give whole number results at the anchor points, but were otherwise uniformly lower than N_2 scores. His published results appeared to show a relatively smooth curve of scores as V_1 increases, and he confidently concluded that:

> [M]y index outperforms [the effective number of parties] as an operationalization of the variable number and size of parties . . . [because it] behaves better in relation to the size of the largest party and the gap between the two largest parties.

> (1991: 1383, 1391)

These claims have not been widely endorsed, nor has the Molinar index been much adopted by other authors. M scores are slightly trickier to calculate; it is hard to work out the rationale of the calculations involved; and almost no other commentators have seen it as an intuitively useful measure. Lijphart (1994: 69–70) discusses a hypothetical situation where two parties start off receiving half the votes each, and then move to vote shares of 50, 25 and 25. He points out that the Molinar score goes *down* from 2 to 1.89, going against the grain of what he would intuitively expect.

To analyse how this conundrum comes about we apply the same approach as above, computing the Molinar index (M) scores under maximum and minimum fragmentation conditions, and showing how these scores vary with the level of V_1 and the number of relevant parties competing. Figure A1 replicates Figure 2 for the M index, and also shows the $1/V_1$ line to facilitate comparison. It is apparent that the spaces for the number of party scores under the Molinar index are patterned in a complex way. In particular, Figure A1 shows that the spaces for scores with up to eight parties 'cross-over' the $1/V_1$ line at various points when V_1 levels are between 30 and 45 percent. If V_1 is greater than the cross-over point, then three things happen: the available space for scores lies *below* the $1/V_1$ line; the minimum fragmentation line is at the *top* of each space; and the maximum fragmentation line is at the *bottom* of the space – all of which is the exact opposite of the situation for N_2 in Figures 2 and 3. However, if the largest party's support is below the cross-over point in Figure A1, then the available space for scores lies *above* the $1/V_1$ line; the minimum fragmentation line is now at the bottom or left of the relevant party space; and the maximum fragmentation line is at the top – all of which are the same as with the N_2 index.

Again the space for scores when there are x relevant parties includes all the spaces for scores when the number of parties is below x, so the space for 8 parties includes also those for 7, 6, 5, 4 and 3 parties. Where the spaces reach the V_1 line above the cross-over point the shape of the spaces also shows a pattern of 'bobbling' and kinks around the anchor points quite like those with the effective party measure. But these
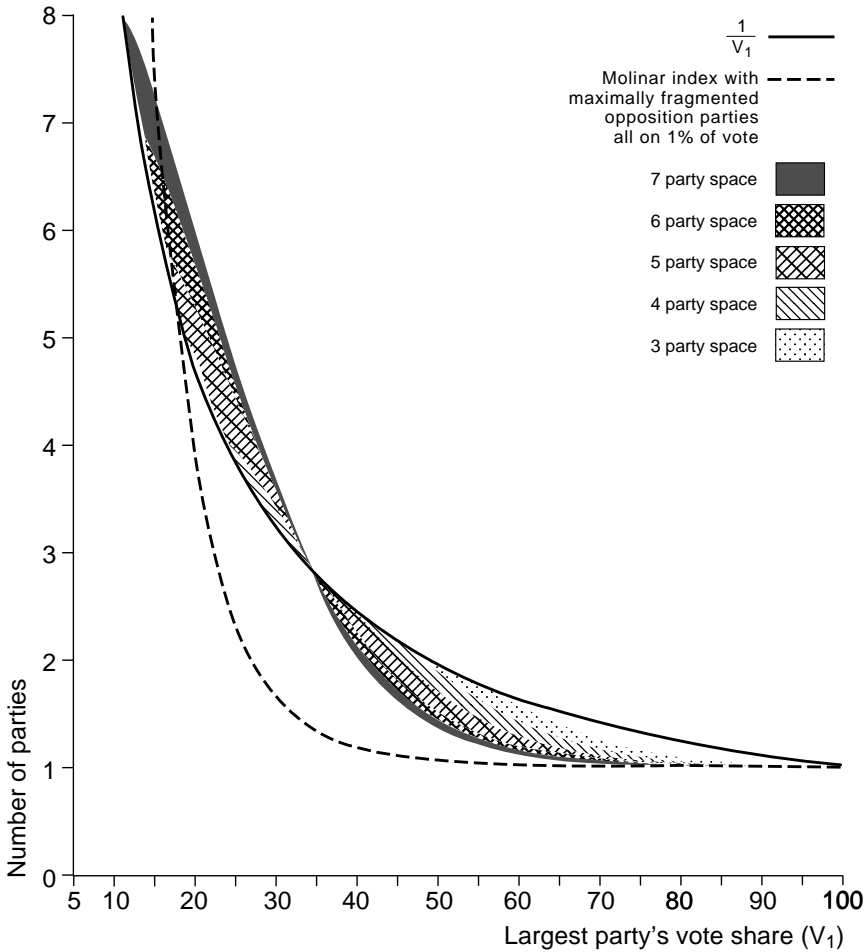
Figure A1. How the shape of the space for the Molinar index varies with the number of relevant parties

detailed effects are hardly visible in Figure A1 because the M index compresses the spaces for party scores more than any of the indices reviewed above, especially in the vicinity of the cross-over points. At very low levels of V_1 (below 16 percent) the additional spaces for parties hug a nearly vertical $1/V_1$ line, so the effect of fragmentation on the M scores is hard to show visually.

The key conclusion from Figure A1 is that when the largest party has high or medium support the M index works in the opposite way to not only N_2 but to all other counting and weighting methods, scoring situations with more relevant (or observable) parties as having fewer weighted parties. By contrast, when V_1 levels fall below the cross-over points in Figure A1, fragmented oppositions are scored in a conventional manner as having more weighted parties than those with a concentrated opposition. A simple numerical illustration for six relevant parties may help to show how the M index behaves differently from the N_2 index:

Fragmentation	V_1	V_2	V_3	V_4	V_5	V_6	M score	N_2 score
Minimum	50	46	1	1	1	1	1.95	2.16
Maximum	50	10	10	10	10	10	1.56	3.33
Minimum	25	25	25	23	1	1	4.08	4.16
Maximum	25	15	15	15	15	15	4.67	5.71

Under the Molinar index the crucial cross-over point (which has such a big impact on the index's performance) moves gradually leftwards as V_1 levels fall. With three relevant parties the cross-over occurs with V_1 at 42 percent, and moves to 40 percent with 4 parties, 38 percent with 5 parties, 36 percent with 6 parties and so on. We dramatize this effect in Figure A1 by showing what happens to the M score as V_1 falls, if the opposition vote is always fragmented into as many parties as possible given that V_1 level (each opposition party with 1 percent support). At this maximal level of party fragmentation, the M line is almost flat (at a score of 1 party) until V_1 falls below 50 percent. It then rises slowly at first and then faster until it is nearly vertical, crossing over the $1/V_1$ line at around 16 percent. If there is no floor on the minimum size of observable parties, and so the number of opposition parties can reach infinity, then the M index's maximum fragmentation limit remains at 1 for all levels of V_1.

We can assess Molinar's approach in terms of three distinct features.

(i) The cross-over points mean that M behaves in different ways depending on the level of V_1. Molinar's argument in defence of this pattern in some subterranean way seems to want to incorporate into the number of parties score a measure or prediction of the *governing* or *coalitional* power of the largest party. But there are already many ways of measuring V_1's power directly, such as the Shapley-Shubik or Banzhaf indices. Alternatively, where actors' preferences are known, an appropriate game-theoretic measure could incorporate spatial information. The case for not mixing up a number of parties index with an attempt to capture V_1's power is simply Isaiah Berlin's dictum that: 'Everything is what is, and not another thing.' Determining the number of weighted parties competing for an electorate's votes or winning seats in a legislature is not the same thing as assessing the largest party's power consequent on these outcomes. These phenomena need to be independently measured on different dimensions, not conflated.

(ii) The cross-over points shift slowly leftwards in Figure A1 as the number of relevant parties increases. How does this feature sit with Molinar's claim that his index has greater 'intuitive' appeal than N_2? Perhaps this idea is already latent in our ordinary ways of thinking about these issues, but if so it must be in some very un-obvious way. Perhaps the varying Molinar cross-over points go unrecorded in our intuitive thinking but are still in fact important 'magic numbers', of which we need to take cognisance in future research. The alternative possibility is that the cross-over points are arbitrary mathematical quirks, with no counterpart in our conventional thinking about these issues, nor any wider importance outside the M index itself. We incline to this view.

(iii) Lijphart concluded about the Molinar index that: 'the alternative indices to the effective number of parties do not differ from it in essentials' (1994: 70). But consider Figure A2, which shows a situation with seven relevant parties under both the N_2 and the M indices. The space for the effective number of parties scores is
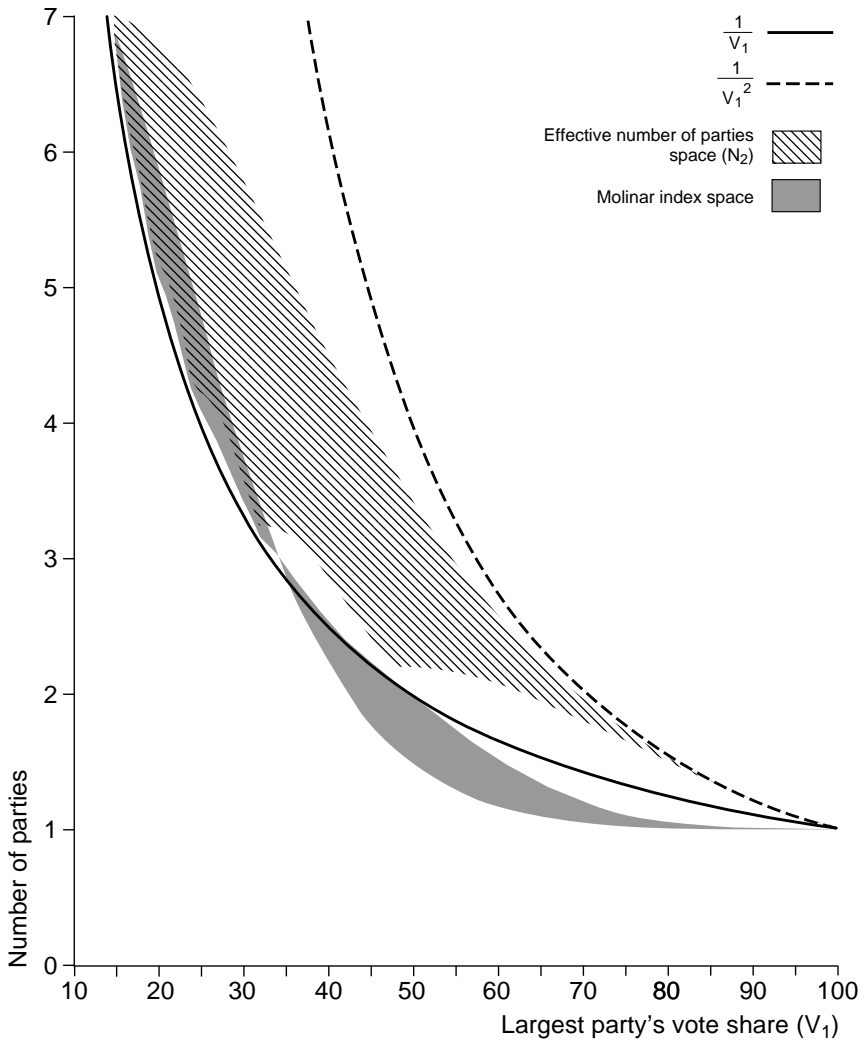
Figure A2. Comparing areas for the effective number of parties index (N_2) and the Molinar index with seven relevant parties

shown using hatched lines, while the space for the M scores is shown shaded grey. The two indices offer the same numerical estimate at the terminal anchor point (and when $V_1 = 100$ percent). But otherwise there is no overlap of the two spaces. (The apparent visual overlap of the two areas in Figure A2 might suggest that in some situations the indices could produce the same score, but in fact this is not the case.)[1] Indeed, these spaces could hardly be more different. The figure also demonstrates the oddness of both indices' spaces – both the extended 'bat-wing' shape of N_2 with its prominent kinks in the minimum fragmentation boundary, and the narrow 'two-wing' shape of the Molinar index (with some smaller kinks in its minimum

fragmentation boundary at lower V_1 levels). Both areas contrast strongly with the smooth areas defined by the $1/V_1$ and $1/V_1^2$ curves. Can either N_2 or M make realistic claims to fit with our 'intuitive' judgements of the weighted number of parties – which surely do not extend to such precise and yet hard-to-predict effects?

These considerations all suggest to us that when the Molinar index was initially proposed its author did not fully appreciate that it behaves in such a complex way, and nor have subsequent commentators. In our view, the index should be dropped forthwith from its residual place in the political science toolkit.

Notes

We are deeply indebted to Mina Moshkeri of the Design Unit at the London School of Economics and Political Science for expertly producing the figures included here. We also thank Keith Dowding, Matthew Mulford and Gordon Smith of LSE, and Kennedy Stewart of Simon Fraser University, Vancouver for comments on an earlier version of this article. We are grateful to David Farrell and anonymous referees of the journal for constructive criticism and comment.

1 If we write a for the sum of all parties' squared vote shares, and b for the largest party's squared vote share, both of which can vary between 0 and 1, then the identity condition for N_2 and M is: $1/a = 1 + [(a − b/a)*1/a]$.

Rearranging we get: $1 + [(a − b)/a2] − 1/a = 0$. Multiplying through by a we get: $1 + \{[a(a − b) − a2]/a3\} = 0$, which simplifies to: $1 − b/a2 = 0$ or: $b/a2 = 1$.

Only if $a = b = 1$ can this condition be met, which would mean that the sum of all parties' squared vote shares equals the largest party's squared vote share. This can only occur if either there are no opposition parties or there is an infinitely large number of opposition parties (each with infinitely small vote shares).

References

Colomer, J. (2001) *Political Institutions: Democracy and Social Choice*. Oxford: Oxford University Press.

Cox, G. W. and M. S. Shugart (1991) 'Comment on Gallagher's "Proportionality, Disproportionality and Electoral Systems"', *Electoral Studies* 10: 348–52.

Dunleavy, P. (1996) 'Political Behaviour: Institutional and Experiential Approaches', in R. E. Goodin and H.-D. Klingemann (eds) *A New Handbook of Political Science*, pp. 276–93. Oxford: Oxford University Press.

Dunleavy, P. and H. Margetts (1994) 'The Experiential Approach to Auditing Democracy', in D. Beetham (ed.) *Defining and Measuring Democracy*, pp. 155–81. London: Sage.

Gallagher, M. (1991) 'Proportionality, Disproportionality and Electoral Systems', *Electoral Studies* 10: 33–51.

Laakso, M. and R. Taagepera (1979) ' " Effective" Number of Parties: A Measure with Application to West Europe', *Comparative Political Studies* 12: 3–27.

Lane, J. E. and S. O. Ersson (1987) *Politics and Society in Western Europe*. London: Sage.

Lijphart, A. (1984) *Democracies: Patterns of Majoritarian and Consensus Government in Twenty-One Countries*. New Haven: Yale University Press.

Lijphart, A. (1994) *Electoral Systems and Party Systems*. Oxford: Oxford University Press.

Lijphart, A. (1999) *Patterns of Democracy*. New Haven, CT: Yale University Press.

Lindblom, C. and D. Cohen (1979) *Useable Knowledge*. New Haven, CT: Yale University Press.

Mair, P. (1987) 'Types of Party Systems', in V. Bogdanor (ed.) *The Blackwell Encyclopaedia of Political Institutions*, pp. 420–2. Oxford: Blackwell.

Mair, P. (1997) *Party System Change: Approaches and Interpretations*. Oxford: Oxford University Press.

Mayer, L. C. (1980a) 'Party Systems and Cabinet Stability', in P. Merkl (ed.) Western *European Party Systems: Trends and Prospects*, pp. 335–47. New York: Free Press.

Mayer, L. C. (1980b) 'A Note on the Aggregation of Party Systems', in P. Merkl (ed.) *Western European Party Systems: Trends and Prospects*, pp. 516–20. New York: Free Press.

Molinar, J. (1991) 'Counting the Number of Parties: An Alternative Index', *American Political Science Review* 85: 1383–91.

Rae, D. (1967) *The Political Consequences of Electoral Laws*. New Haven, CT: Yale University Press.

Rokkan, S. (1970) *Citizens, Elections, Parties: Approaches to the Study of Comparative Development*. Oslo: Universitetsforlaget.

Taagepera, R. (1997) 'Effective Number of Parties for Incomplete Data', *Electoral Studies* 16: 145–51.

Taagepera, R. (1999) 'The Number of Parties as a Function of Heterogeneity and Electoral System', *Comparative Political Studies* 32(5): 531–48.

Taagepera, R. and M. Laakso (1980) 'Proportionality Profiles of West European Systems', *European Journal of Political Research* 8: 423–36.

Taagepera, R. and M. S. Shugart (1989) *Seats and Votes: The Effects and Determinants of Electoral Systems*. New Haven, CT: Yale University Press.

Ware, A. (1995) *Political Parties and Party Systems*. Oxford: Oxford University Press.

Wildgen, J. K. (1971) 'The Measurement of Hyperfractionalization', *Comparative Political Studies* 4: 233–43.

PATRICK DUNLEAVY is Professor of Political Science and Public Policy and Chair of the LSE Public Policy Group at the London School of Economics and Political Science. He co-edited *Developments in British Politics 7* (2003) and previous volumes in the same series, and wrote *Democracy, Bureaucracy and Public Choice* (1991). Recent articles have covered coalition formation and change in the UK's electoral systems.

ADDRESS: Department of Government, London School of Economics, Houghton Street, London WC2A 2AE, UK. [email: p.dunleavy@lse.ac.uk]

FRANÇOISE BOUCEK is Researcher in the LSE Public Policy Group at the London School of Economics and Political Science, working on focus group research. In her

LSE doctorate, completed in 2002, she studied factionalism in long-lived dominant parties and she is working up papers from the thesis for journal publication. She has published papers on change in dominant party systems and on trends in French politics.
ADDRESS: LSE Public Policy Group, London School of Economics, Houghton Street, London WC2A 2AE, UK. [email: f.a.boucek@ lse.ac.k]