

UC Irvine

UC Irvine Previously Published Works

Title

Mapping the Indices of Seats-Votes Disproportionality and Inter-Election Volatility

Permalink

<https://escholarship.org/uc/item/0m9912ff>

Journal

Party Politics, 9(6)

ISSN

1354-0688

Authors

Taagepera, Rein
Grofman, Bernard

Publication Date

2003-11-01

DOI

10.1177/13540688030096001

Peer reviewed

MAPPING THE INDICES OF SEATS–VOTES DISPROPORTIONALITY AND INTER-ELECTION VOLATILITY

Rein Taagepera and Bernard Grofman

ABSTRACT

Measures of electoral system disproportionality and of party system volatility (as well as malapportionment and vote splitting) present similar statistical issues in terms of deciding what index is most appropriate, but it is not common to view indices of disproportionality and volatility as serving similar ends. Making use of 12 different criteria, we evaluate 19 indices that have been previously proposed as measures of either disproportionality of electoral seats–votes results or over-time volatility of party vote (or seat) shares. We suggest that, on balance, Gallagher's (1991) index, which has achieved increasing acceptance in the seats–votes literature on disproportionality (see esp. Lijphart, 1994) offers the most desirable combination of features, although the advantages it offers over the Loosemore–Hanby index are not large and are debatable. We also find that Dalton's principle of transfers presents an ambiguity when one party has a larger number of excess seats, while another has a larger proportion of them.

KEY WORDS ■ Dalton's principle of transfers ■ malapportionment ■ proportionality ■ volatility of party shares ■ vote splitting

The Objective: Mapping the Measures of Discordance

Deviation from proportional representation (PR) means the difference in party vote shares and party seat shares in some given election (v_i versus s_i). Volatility of votes means the difference in party vote shares (v_j) from one election to another (v_{it} versus v_{jt}); similarly, volatility of seats means the difference in party seat shares (s_j) from one election to another (s_{it} versus s_{jt}). How do we best measure the proportionality/disproportionality of the translation of party vote shares into party seat shares in a given election?

How do we best measure the stability/volatility of party vote shares (or party seat shares) across elections?

The first question has generated a number of studies dealing with measurement issues and empirical applications, and well over a dozen different measures have been proposed (see especially the excellent review in Monroe, 1994 and the discussion below). For the second question, one particular measure, the absolute difference between party vote shares (or seat shares) across two time periods, has been widely accepted (see, e.g., Bartolini and Mair, 1990: Ch. 1). Both questions raise important normative and empirical issues. Disproportionality is key to ideas of electoral fairness, and volatility is a key element in understanding party system stability and realignment. The choice of measure may affect the conclusions we reach. For example, by one index of disproportionality, deviation from perfect proportionality may seem reduced, while by another it seems increased.

The literatures dealing with each of these two questions have been distinct. This isolation of the two literatures might appear reasonable in that (1) the first question involves analysis of a single election while the second requires us to look at two or more elections; and (2) the first question involves comparisons between two distinct variables, seat shares and vote shares, while the second looks only at one or the other of these two variables. In our view, however, it is useful to consider the answers to these two questions in terms of a unified perspective. Any of the proposed measures of disproportionality can also be applied, with a straightforward change in notation, to measure volatility. And, conversely, indeed, the most common measure of volatility is isomorphic to what had been, until quite recently, the most commonly used disproportionality measure, the Index of Distortion, D (Loosemore and Hanby, 1971).

As Monroe (1994) has stressed, *malapportionment* is a related aspect. While the political content is different, the dilemmas are similar mathematically. A fourth related aspect should also be pointed out, that of *vote splitting* (and here we are not aware that a formal connection has been discussed by any earlier author). If voters have two ballots (president and assembly, two different chambers, a district and a nationwide vote for the same chamber, etc.), then the measurement of the degree of disagreement between the two votes (v_A versus v_B), i.e. vote splitting or split ticket voting, presents conceptually the same types of measurement issues as volatility and deviation from PR. There is one difference. In volatility one election comes first, in deviation from PR votes come before seats, and in malapportionment voters come before seats, thus representing a comparison point for what comes later. In contrast, neither of the simultaneous two ballots involved in vote splitting may come 'first': they may be conceptually on an equal footing.¹

All four aspects force us to deal with *discordance between two sets of figures that conceivably can be equal* and maybe ideally should, if one imposes norms such as perfect PR, perfect voter loyalty, perfectly

proportional apportionment, or perfect consistency of voting.² In the following, we focus on deviation from PR and volatility, but the implications for malapportionment and vote splitting should also be kept in mind.

Our purpose is threefold:

- First, we give a systematic overview of the indices of disproportionality that have been proposed in the literature, and we show how these indices might also be applied to measure volatility.³
- Second, building on work by Monroe (1994), we identify 12 different criteria which we argue it is desirable for any index of disproportionality or volatility to satisfy. Moreover, for all of the indices of disproportionality or volatility that have been seriously offered in the literature, we specify which of our criteria are satisfied by that index. In so doing we note an ambiguity in Dalton's principle of transfers when it is applied to parties of unequal sizes (rather than individuals).
- Third, after determining that the measure proposed by Loosemore and Hanby (1971), the analog to which is commonly used to measure volatility, has a number of desirable properties, we argue that the index in Gallagher (1991), whose use for measuring seats-votes disproportionality has recently been argued for by Lijphart (1994), is (marginally) to be preferred.

Nineteen Previously Proposed Indices

Notation

For both volatility and deviation from PR, we are talking about discordance between two sets of figures that conceivably can be equal and maybe ideally should. It is deviation from some conceptual baseline: for the latter the baseline is that seat and vote shares are equal, for the former it is that party vote (or seat) shares in two elections remain unchanged. We want a notation that can be applied to all these problems. For volatility, we use x_i instead of v_{it} or s_{it} and y_i instead of $v_{it'}$ or $s_{it'}$. Similarly, for proportionality we use x_i instead of v_i and y_i instead of s_i . (For vote splitting, either ballot might be taken as x .)

This notation has the additional advantage that it makes more transparent the relationships between measures of disproportionality or volatility and standard statistical concepts such as variance, sum of squares, entropy and chi-square. In general, x and y are taken as fractional values out of 1 (rather than 100%). In numerical examples, however, we convert them into percentages. To avoid the need to keep saying 'proportionality/volatility, etc.' we discuss our findings primarily in terms of the concept of disproportionality.

Use the following notation:

c = a normalizing constant – e.g., for the Loosemore and Hanby (1971) index of distortion, D , and for Gallagher's (1991) index, Gh , $c = \frac{1}{2}$. Because these constants do not really affect the indices, except as to bounds, we have shown them generically.

P = the total number of parties and independents running.

n = the number of degrees of freedom ($= P - 1$).

V and S = the total number of votes and seats, respectively.

$A_i = y_i/x_i$. This concept of the 'advantage ratio' of a party is due to Taagepera and Shugart (1989). Perfect proportionality for a party corresponds to $A_i = 1$.

N = the 'effective number' of parties (Laakso and Taagepera, 1979), which can be calculated in terms of either votes or seats (or earlier or later elections).

Seven Pre-1990 Indices of Deviation from Proportional Representation

A huge number of indices (and entire families of indices) are conceivable in principle and have been pointed out by Monroe (1994), Pennisi (1998) and Grilli di Cortona et al. (1999: 85–107), among others. We have restricted our scope here to the ones that have been explicitly recommended or used by some authors. We distinguish between the pre-1990 and post-1990 indices to highlight the fact that, instead of converging toward the use of one or a few indices, the field of alternatives has lately expanded at a faster rate than ever.

In the order of their appearance in the electoral literature, the following indices were considered prior to 1990.

$I = (1/P)\sum y_i - x_i$. When Rae (1967) first proposed and calculated this index, parties with less than 0.5 percent votes were omitted.

$D = c\sum |y_i - x_i|$, where $c = \frac{1}{2}$. The usual citation to this 'index of distortion' is Loosemore and Hanby (1971). However, as an 'index of dissimilarity' it harks back at least to Duncan and Duncan (1955). Until recently, it predominated as a measure of deviation from PR, and it has also been used for volatility: from Przeworski (1975), who called it 'deinstitutionalization', to Coppedge (1996). In volatility studies, it is often called the Pedersen index.

$I' = c[|y_1 - x_1| + |y_2 - x_2|]$, where $c = \frac{1}{2}$. It involves the two largest x_i only. Lijphart (1984: 163) proposed this index, which can be viewed as a truncation of either I or D . While Lijphart himself (1994) has given up on its use, it continues to be used by some other authors. For more than two parties with a non-zero deviation from PR, $I' < D$. However, I' can be smaller or larger than I . The mathematical equivalent of I' in volatility measurement is 'swing', that usually takes into account only the changes for the largest two parties. It harks back at least to David Butler's study of the British 1951 election.

χ^2 (chi-square) for the expression $\sum (y_i - x_i)^2/x_i$. Nagel (1984) is the first

author whom we have identified who suggests applying chi-square to disproportionality data. This approach is also discussed in Gallagher (1991). Pennisi (1998) uses the equivalent of χ^2 multiplied by S^2/V , and also the same expression with the roles of seats and votes (x and y) reversed ('Equal Proportions'). The latter tends to infinity when a party with some votes obtains zero seats. We will not consider these forms separately from χ^2 .

$Gr = (1/N_x)\sum |y_i - x_i| = \sum x_i^2 \sum |y_i - x_i| = 2D/N_x$. In the mid-1980s, Grofman (personal communication, cited in Lijphart, 1994: 61) proposed this index as a way of creating an index of 'average' rather than 'total' disproportionality. In this respect it is similar to I . In general, $I < Gr < D$, and $I' < Gr < D$, but exceptions can be found.

$T\&S = (N_x - N_y)/N_x = 1 - \sum (x_i^2)/\sum (y_i^2)$. Taagepera and Shugart (1989: 273) call this 'relative reduction in the number of parties', and use r to denote it. However, to avoid confusion with the Pearson correlation coefficient, which is also commonly denoted r , we label it $T\&S$. In general, $T\&S > D$, but exceptions can occur, and $T\&S$ can assume small negative values.

Gini index of inequality. While they do not provide any numerical calculations, Taagepera and Shugart (1989: 263) suggest that the well-known Gini index of inequality could be applied as a measure of disproportionality. For definition of and discussions of this index, see, e.g., Monroe (1994), White (1986).

Twelve Indices Proposed Since 1990

$Gh = [c\sum (y_i - x_i)^2]^{0.5}$, where $c = 1/2$. This is what Gallagher (1991) refers to as his 'least square measure'. We prefer to call it simply Gallagher's measure, since the 'least square' is somewhat of a misnomer because no minimization of squares is involved.⁴ Apart from the factor c , Gh is the geometric formula for the length of a line segment in terms of the coordinates of its endpoints in a P -dimensional space. We might also think of x_i as the expected value for y_i , if perfect proportionality obtained (or if volatility was zero); then, apart from the factor c , Gh is also recognizable as an element in the formula for the correlation coefficient.⁵ Supported by its use in Lijphart (1994), Gh has rapidly become the major competitor to D in the electoral studies literature. It can be shown that $Gh < D$ when more than two components have non-zero deviation from PR.

a = the slope of the regression line of y versus x ($y = ax + b$). Cox and Shugart (1991) proposed this measure. For perfect equality, $a = 1$. Larger slopes indicate large-party advantage, while $a < 1$ indicates a small-party advantage. An advantage of a over other indices of disproportionality is that it indicates the directionality of the imbalance. We can express a in terms of y and x by making use of well-known results about the slope of the best-fitting regression line of y on x .⁶ Thus $a = \sum (y_i - \bar{y}) (x_i - \bar{x}) / \sum (x_i - \bar{x})^2$.

$D_p = [\sum (y_i - x_i)^2 / (1 + \sum x_i^2)]^{0.5}$. This index is found in Monroe (1994), and it is visibly equal to $Gh[2/(1 + \sum x_i^2)]^{0.5} = Gh[2/(1 + 1/Nv)]^{0.5}$. Unless there is only one component ($x_i = 1$), $D_p > Gh$.

$\max A = \text{maximum } y_i/x_i$. The notion of looking at the maximum advantage ratio as a measure of disproportionality is due to Gallagher (1991). A problem with this measure is that its value would often be based on the outcomes for a very small party (or an independent), but it could be specified that parties with less than 10 percent votes are to be excluded (except if such a party is the largest one).

$\beta = \ln(x_1/x_2)/\ln(y_1/y_2)$. Galeotti (1994) proposed this index. Like I' , this index is based on values for the two largest x_i only.

$L = \max |y_i - x_i|$. This measure is the analog of $\max A$, when replacing ratio by difference. It is found in Lijphart (1994: 62). Although Lijphart prefers Gh to L as a measure of disproportionality, he remarks that L correlates almost perfectly with Gh ($r = 0.99$). Rein Taagepera, in unpublished work, has shown that, for volatility data for a set of 25 post-World War II (PWII) democracies, the mean values of L and Gh are virtually identical, while at the individual country level, $0.77Gh < L < 1.23Gh$. L is usually based on the over-representation of one of the largest parties, but occasionally a catastrophic loss by a third party can exceed the individual gain of either of the two largest parties.

$PWI = \sum x_i/y_i - x_i$. Li (1995) offers this as a weighting scheme that deals with the problem that D supposedly over-represents deviations for small parties. PWI is intended to be measured in terms of y_i and x_i expressed as fractional shares. Like D_p , and unlike the other measures we have considered so far, PWI suffers from the disadvantage that it depends on whether fractional shares or percent shares are used.⁷

Chi-square-based probability = $1 - p[t < \chi^2(n)]$, where χ^2 is based on $\sum (y_i - x_i)^2/x_i$ and n is the number of degrees of freedom ($= P - 1$). Mudambi (1997) offers this measure as a way of assuring values between 0 and 1 that reflect the likelihood that deviations from proportionality of the observed magnitude could have occurred by chance. The Mudambi measure is simply the probability associated with the chi-square calculation suggested by Nagel (1984).

$L_1 = S/V \sum |y_i - x_i|/x_i = S/V \sum |A_i - 1|$ in terms of advantage ratios. This is, in our notation, the 'L₁-norm' used by Pennisi (1998). Where Li (1995) multiplies by x , Pennisi (1998) divides by x , thus enhancing the impact of small parties even more than is the case with D .

$L_2 = (S/V)^2 \sum (y_i - x_i)^2/x_i^2 = (S/V)^2 \sum (A_i - 1)^2$ in terms of advantage ratios. This is, in our notation, the 'L₂-norm' used by Pennisi (1998). Where χ^2 divides by x , this form divides by x^2 , thus enhancing the impact of small parties.

$L_\infty = S/V \max |y_i - x_i|/x_i = S/V \max |A_i - 1|$. This is, in our notation, the 'L_∞-norm' used by Pennisi (1998). Like $\max A$ and L , it maximizes an expression in y_i and x_i , in this case the one used in L_1 . Unless small parties

(and independents) are excluded, L_∞ is highly likely to be determined by a party that wins no seats (so that $|A_i - 1| = 1$) or has a huge advantage ($A_i > 2$) that only a tiny party can have.

Entropy = $\sum y_i \ln(y_i/x_i)$. Entropy is of course a well-known quantity, but we are not aware of it having been used for disproportionality prior to Pennisi (1998).

The Criteria an Index Should Satisfy

We now propose that it would be advantageous for indices of volatility and deviation from PR to satisfy the following 12 criteria. The first seven are of a theoretical nature. Among these, we discuss the principle of transfers separately, because we have found a little-known paradox. The last five criteria are of considerable practical importance in terms of ease of calculations and unambiguity of results. Our criteria overlap considerably but not fully with those presented by Monroe (1994).

Theory-Inspired Criteria

An index of volatility and deviation from PR should satisfy the following theoretically grounded criteria.

1. Is *informationally complete*: makes use of the x_i and y_i data for all parties. Indices that use only the two largest parties (Lijphart's I' , Galeotti's β) or the largest difference (L), ratio ($\max A$) or a combination of them (L_∞) do not satisfy this criterion.

2. Uses the data *uniformly for all parties*, meaning for instance, that no special role is given to the largest party (x_1, y_1) or the two largest. Most indices that fail completeness also fail on uniformity, but maximizing expressions ($\max A, L$ and L_∞) pass, since every party has an equal chance, in principle, to have the largest value.

3. Uses x_i and y_i *symmetrically*. It might be argued that this is a superfluous requirement for volatility or deviation from PR because, in the notation we have used, x and y are distinguished from one another, and, more importantly, for seats-votes relationships, it would seem reasonable that x (votes) is the independent and y (seats) be the dependent variable. However, when comparing the vote constellations at two simultaneous elections (vote splitting), it should not matter which one is placed first. This criterion is contravened by indices that multiply or divide by x_i without doing the same with y_i ($PWI, \chi^2, 1 - p[t < \chi^2(n)], L_1, L_2, L_\infty$); those which incorporate N_x but not N_y , or conversely (Gr, D_p); those which use division of y by x (or conversely) that thus asymmetrically treat x and y ($\max A$, Galeotti's β , entropy) or which use division of N_y by N_x (or conversely) and thus asymmetrically treat N_y and N_x ($T\&S$). The regression slope a also violates this criterion, because the slope of y versus x is not the same as the

slope of x versus y (unless $R^2 = 1$ or the x and y variables are both standardized).

4. Varies only *between 0 and 1 (or 100 percent)*. This is contravened by L_1 , L_2 , L_∞ , $\max A$ and β (which can take values much larger than 1) and also by $T\&S$ and the regression slope a (which both can become negative).

5. Has *value 0 (or 0 percent)*, if $y_i = x_i$ for all i , to reflect perfect concordance between the two sets (y and x).⁸ This is contravened by maximum advantage ratio ($\max A$), regression slope a and Galeotti's β .

6. Has *value 1 (or 100 percent)*, if $y_i = 0$ for all $x_i > 0$ and $x_i = 0$ for all $y_i > 0$, because this is the utmost or perfect disproportionality or lack of concordance. Assume that a prince allows elections, but then appoints an assembly of people who did not run, declaring 60 percent of them to represent the nobility and 40 percent the clergy:

x	25	25	25	25	0	0
y	0	0	0	0	60	40

In terms of volatility, this is the situation where all previous parties vanish and are replaced by brand new parties that cannot be traced back to the previous ones.

I , I' , L , Gr , PWI , Gh , D_p and $T\&S$ yield values below 100 percent, and β is indeterminate, while *chi-square* and $\max A$ become indeterminate when one vote share is zero, and the regression slope a is negative. L_1 , L_2 , L_∞ are all much larger than 1 (100 percent). Entropy becomes infinite for volatility when a party with no previous votes starts obtaining some.

7. Satisfies *Dalton's principle of transfers* (see Monroe, 1994): when a seat is transferred from a richer component to a poorer one, the disproportionality index should decrease. Conversely, transfer from a poorer to a richer component (or among initially equal components) should increase disproportionality. The *strong principle of transfers* requires that the index change in the appropriate direction. The *weak principle* is satisfied, as long as there is no change in the wrong direction. This looks like a reasonable and straightforward requirement. However, we run into a little-noticed issue of what do 'richer' and 'poorer' mean, to be discussed next.

The Transfers Dilemma: Differences or Ratios?

Consider the following example, for a 100-seat assembly. The differences ($\Delta_i = y_i - x_i$) and the advantage ratios ($A_i = y_i/x_i$) are shown along with x and y . The values of three frequently used indices are also shown.

7a	x	50	40	10	$D = 15.0\%$
	y	60	25	15	$Gh = 13.23\%$
	Δ	+10	-15	+5	$Gini = 16.50\%$
	A	1.20	0.625	1.50	

Among the two components that are overpaid (compared to PR), which one is the 'richer'? If we go by the number of extra seats, it's the largest one. But the third-largest party is overpaid by 50 percent, while the largest one is overpaid by only 20 percent. Now suppose that one seat is transferred from the largest to the third-largest party:

7b	<i>x</i>	50	40	10	<i>D</i> = 15.0%
	<i>y</i>	59	25	16	<i>Gh</i> = 13.08%
	Δ	+9	-15	+6	<i>Gini</i> = 17.10%
	<i>A</i>	1.18	0.625	1.60	

D remains the same, reflecting its well-known failure to follow the *strong* principle of transfers when both components involved are overpaid (or both are underpaid). *Gh* decreases, because the largest difference has decreased. But *Gini* actually increases, because the largest advantage ratio has increased. We may prefer to correct for the excessive advantage ratio by transferring a seat from the third-largest to the largest party (in 7a):

7c	<i>x</i>	50	40	10	<i>D</i> = 15.0%
	<i>y</i>	61	25	14	<i>Gh</i> = 13.45%
	Δ	+11	-15	+4	<i>Gini</i> = 15.90%
	<i>A</i>	1.22	0.625	1.40	

Now *D* still remains the same, but *Gh* increases, while *Gini* decreases. Generalizing: when one among the overpaid parties has a larger difference and the other a larger ratio, *Gh* and *Gini* vary in opposite directions. If one decides that 'richer' means having a larger excess of seats, then *Gh* follows the strong principle of transfers and *Gini* fails even the weak one. Conversely, if 'richer' means having a larger excess proportion of seats, then *Gini* follows the strong principle of transfers and *Gh* fails even the weak one. Meanwhile, *D* always follows the weak principle and fails the strong one, thus being more neutral. In the numerous cases where the same party (often the largest) has both the largest difference and also the largest ratio, of course, *Gh* and *Gini* pull in the same direction and have an advantage over *D*.

Should ratio or difference considerations prevail when testing for the principle of transfers? Political scientists probably would opt for the difference. It's the number of extra seats for a large party that matters in ease of cabinet formation, rather than a large relative excess for a small party that still remains small. Yet the mind baulks at adding seats to a party that is already 50 percent overpaid and calling it a reduction in disproportionality!⁹ Maybe this is the crucial difference between the notions of 'deviation from proportionality' (stressing the difference) and 'inequality' (stressing the ratio).

If difference is taken as the criterion of being rich, all indices involving squares of differences (*Gh*, D_p , χ^2 , $1 - p$, L_2) do satisfy the strong principle

of transfers, and *Gini* and *entropy* can fail even the weak principle. Indices based on absolute values (*D*, *I*, *I'*, *Gr*, *PWI*, *L*₁, *L*_∞) fail to change when seats are shifted among over-represented parties, and thus fail the strong principle but not the weak one. *L* does change in the examples above, but fails to do so when shares are reshuffled without affecting the largest difference. The same is true of β when the two largest components are not involved, and of *max A* when the largest ratio is not involved. But neither do these indices decrease, thus satisfying the weak principle. The regression slope coefficient, *a*, satisfies the strong criterion in the examples above and is given the benefit of the doubt in the general case, and the same applies to *T&S*.

Practical Criteria

In terms of ease of calculations and unambiguity of results indices should satisfy the following.

8. Does *not include P*, the number of parties, which is often hard to define in the presence of small parties or independents and, in any case, is quite sensitive to the number of truly minor parties or independents. Should a party (or an independent) that receives only one vote be included? Or one that runs but receives 0 votes (with independent candidates this has happened!)? This criterion is contravened by *I* and $1 - p[t < \chi^2(n)]$. Of course, arbitrary cut-offs can be introduced, but arbitrary cut-off points are to be avoided because this would contravene Criterion 2 (using data for all parties uniformly).

9. Is *insensitive to lumping of residuals*. Consider the following two constellations:

9a	x	40	30	20	5	5
	y	50	30	10	5	5

9b	x	40	30	20	9	.1	.1	.1	.1	.1	.1	.1	.1	.1
	y	50	30	10	0	1	1	1	1	1	1	1	1	1

Both could be presented in a data source as:

x	40	30	20	[10]
y	50	30	10	[10]

with the last column lumping minor groups as 'Others'. Such cases are difficult to handle when the index involved is sensitive to small party details. Here, from 9a to 9b, the values of *D*, *I*, *Gr* and *Gini* increase by about a factor of 2, while *chi-square* increases threefold, entropy increases sevenfold, and *max A* increases ninefold. As for $1 - p[t < \chi^2(n)]$, *L*₁, *L*₂ and *L*_∞, they change even more. In contrast, *I'*, *PWI*, *Gh*, *D_p* and β change by a

factor of 1.2 at most.¹⁰ This single example is no proof that the latter indices are always insensitive to lumping residuals, but it demonstrates that the former certainly can fail this criterion. L runs into problems when many small losing parties are lumped and their total surpasses the actual single largest deviation. The regression slope, a , is also sensitive to how residuals are treated.

10. Is *simple to compute*. We count a measure as failing the criterion if lengthy or multistage calculations are needed. We also allow an intermediate category (neither satisfying nor failing the criterion) for indices that use square roots or logarithms.

11. Is *insensitive to shift from fractional to percent shares*. For most indices, percent shares can replace fractional shares, and the formula still works. This is not the case for PWI , which requires fractional shares; using percent shares yields values larger than 100, because they are not really in percent but in 'percent-squared'. This is not a critical requirement, but still a desirable feature, so as to avoid mistakes. There are some intermediary cases. Whenever the effective number of parties is used, its formula is different depending on whether fractional or percent shares are used. This affects Gr . For D_p similar caution is needed. In other cases, percent shares can be used, but the outcome itself is in fractional shares. This is the case for $T\&S$, $Gini$, a , $max A$ and $1 - p$.

12. The input data consist only of vote (and/or seat) shares (x_i and y_i), meaning that it does not depend, for instance, on the total number of votes or seats. L_1 , L_2 and L_∞ fail this criterion, which would also be violated by some malapportionment indices discussed by Monroe (1994).

Relative Advantages of Various Indices

Mapping the Advantages and Disadvantages

We have previously presented 19 indices in order of their appearance in the literature. But this is not a particularly theoretically useful way of grouping indices, although it is useful for purposes of literature review. In Tables 1 to 3 we group indices into three categories. Table 1 lists those indices whose central measure of deviation is based on absolute difference, $|y_i - x_i|$. Table 2 lists those whose central measure of deviation is based on the squares of differences, $(y_i - x_i)^2$. Table 3 lists those which do not fall into either of the above categories.

For each of the 12 criteria we have identified above, Tables 1 to 3 indicate whether each of our 19 indices satisfies (✓) or fails to satisfy (–) the given criterion. The total score is taken as a coarse measure of satisfaction. For this summary tally, criterion satisfaction is counted as 1; failure to satisfy a criterion is counted as 0. We allow for halfway values (0.5) for the criteria of transfer principle (based on difference), simplicity and percentage use. In

Table 1. Evaluation of indices of disproportionality/volatility based on absolute value of difference

Criterion	Indices based on $\sum y_i - x_i $				Other $ y_i - x_i $			
	D	I	I'	Gr	PWI	L	L ₁	L _∞
1. Completeness	✓	✓	–	✓	✓	–	✓	–
2. Uniformity	✓	✓	–	✓	✓	✓	✓	✓
3. Symmetry x-y	✓	✓	✓	–	–	✓	–	–
4. Within (0,1) range	✓	✓	✓	✓	✓	✓	–	–
5. Zero limit	✓	✓	✓	✓	✓	✓	✓	✓
6. 100% limit	✓	–	–	–	–	–	–	–
7. Difference transfer	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
8. No <i>P</i>	✓	–	✓	✓	✓	✓	✓	✓
9. Lumping OK	–	–	✓	–	✓	–	–	–
10. Simplicity	✓	✓	✓	✓	✓	✓	✓	✓
11. Per cent OK	✓	✓	✓	0.5	–	✓	✓	✓
12. Only x and y	✓	✓	✓	✓	✓	✓	–	–
Sum	10.5	8.5	8.5	8	8.5	8.5	6.5	5.5

Tables 2 and 3, the effect of shifting from difference-based to ratio-based principle of transfers is shown separately, at the bottom – all square-based indices are down by one notch, while *Gini*, *max A* and *entropy* are up by

Table 2. Evaluation of indices of disproportionality/volatility based on square of difference

Criterion	Based on $\sum (y_i - x_i)^2$		Other $(y_i - x_i)^2$		
	<i>Gh</i>	<i>D_p</i>	χ^2	$1 - p$	<i>L₂</i>
1. Completeness	✓	✓	✓	✓	✓
2. Uniformity	✓	✓	✓	✓	✓
3. Symmetry x-y	✓	–	–	–	–
4. Within (0,1) range	✓	✓	✓	✓	–
5. Zero limit	✓	✓	✓	✓	✓
6. 100% limit	–	–	–	✓	–
7. Difference transfer	✓	✓	✓	✓	✓
8. No <i>P</i>	✓	✓	✓	–	✓
9. Lumping OK	✓	✓	–	–	–
10. Simplicity	0.5	0.5	✓	–	✓
11. Per cent OK	✓	0.5	✓	0.5	✓
12. Only x and y	✓	✓	✓	✓	–
Sum	10.5	9	9	7.5	7
7a. <i>Ratio</i> transfer	–	–	–	–	–
Sum	9.5	8	8	6.5	6

Table 3. Evaluation of indices of disproportionality/volatility with bases other than absolute value or square of difference

<i>Criterion</i>	<i>Gini</i>	<i>T&S</i>	β	$\max A_i$	<i>a</i>	<i>Entropy</i>
1. Completeness	✓	✓	–	–	✓	✓
2. Uniformity	✓	✓	✓	–	✓	✓
3. Symmetry <i>x-y</i>	✓	–	–	–	–	–
4. Within (0,1) range	✓	–	–	–	–	–
5. Zero limit	✓	✓	–	–	–	✓
6. 100% limit	✓	–	–	–	–	–
7. Difference transfer	–	✓	0.5	–	✓	–
8. No <i>P</i>	✓	✓	✓	✓	✓	✓
9. Lumping OK	–	✓	✓	–	–	–
10. Simplicity	–	✓	0.5	✓	–	0.5
11. Per cent OK	0.5	0.5	✓	0.5	0.5	✓
12. Only <i>x</i> and <i>y</i>	✓	✓	✓	✓	✓	✓
Sum	8.5	8.5	6	3.5	5.5	7.5
7a. <i>Ratio</i> transfer	✓	?	0.5	0.5	?	✓
Sum	9.5	?	6	4	?	8.5

the same amount. We have not investigated the effect on *T&S* and *a* in sufficient detail to pass judgment. (For indices in Table 1 the shift to ratio transfers has no impact.)

No index satisfies all 12 criteria, but *Gh* and *D* come close (10.5), followed at some distance by Monroe’s (1994) *D_p* and by χ^2 (9). We tried to devise an even better index than *Gh* or *D*, but came up short. Needless to say, of course, all criteria do not really carry the same evaluative weight. Thus the reader may discard the ones that seem to her superfluous, weight some others more heavily, and redo the addition. Regardless of how we weight, however, it is apparent from Tables 1 to 3 that some measures carry appreciably more negative baggage than others. It is apparent, too, that *Gh*, *D*, *D_p* and χ^2 will score at or near the top. In view of the dilemma of transfers, *Gini* also should not be counted out of the top group.

Gh is complex to calculate and has the drawback of not reaching 100 percent even at extreme deviation from PR if more than two components have non-zero deviations from PR. Loosemore–Hanby *D* has the advantage of simplicity, but it is sensitive to lumping and fails to follow the strong principle of transfers. As for *Gini*, it is by far the most complex to compute in the presence of many components; but a more serious problem is its sensitivity to the lumping of small components into a residual ‘Other’ category.

How critical are these shortcomings? Moderately involved computation (the weak point of *Gini* and *Gh*) is a minor problem in the age of the computer. Near-total deviations from PR practically never occur; thus this weak aspect of *Gh* might seem secondary. However, it depresses the index

values for nearly all constellations of x and y values. The principle of transfers has some relevance, but it alone would not eliminate D .¹¹ As for the problem of sensitivity to lumped residuals, this is a data problem that occurs all too frequently, and indices sensitive to information hidden in the 'Others' may yield misleading results. All these considerations still favor Gh over D and $Gini$.¹²

Part of the preference for Gh over D has come from the feeling that the values of D are said to be 'too high'. Yet $Gini$ is even higher. It can be shown that $Gh \leq D \leq Gini$ always holds.

What Is a Halfway Deviation from Proportionality?

Our basic difficulty is that we lack a clear rational way to define which constellations have a deviation that must be considered exactly halfway between the ideal extremes of 0 and 100 percent – the extremes which are defined in Criteria 5 and 6. If we had such a definition, the entire scale would be anchored and calibrated within the (0, 1) bounds. We can make an attempt.

Consider the constellations where all vote and seat components are either 50 percent or 0 percent. The rationale is that some of them might be such that deviation from PR would also be self-evidently 50 percent. We can find only three such constellations (short of adding components where both x and y are 0). Consider the two indices with top ratings in Tables 1 and 2 (Gh , D) plus $Gini$.

	case A	case B	case C
x	50 50	50 0 50 0	50 50 0
y	50 50	0 50 0 50	50 0 50
	$Gini = 0\%$	$Gini = 100\%$	$Gini = 75\%$
	$Gh = 0\%$	$Gh = 70.7\%$	$Gh = 50\%$
	$D = 0\%$	$D = 100\%$	$D = 50\%$

Case A involves perfect PR (or zero volatility), so that any self-respecting index should have value 0 (condition 5). All three indices oblige. Case B involves utter deviation: whenever $x = 50$ percent, $y = 0$ percent, and vice versa. Two of the indices yield the expected 100 percent (condition 6), but Gh falls short. Even if both previously existing parties vanish, to be replaced by two brand new ones, volatility according to Gh would be less than 100 percent.

Case C combines one-half of case A with one-half of case B. With one-half at 0 deviation and one-half at 100 percent deviation, could it be conceived as being 'self-evidently' halfway between the extremes? Opinions may vary. However, if one hesitates regarding this constellation, which other constellation could possibly have a stronger intuitive claim to being halfway?

The indices for case C should be at 50 percent. $D = 50$ percent, indeed. Also, $Gh = 50$ percent, but this is not the mean of Gh for cases A and B. $Gini = 75$ percent, which suggests only a relatively minor drop, compared to utter deviation. As we have seen, $Gini$ is sensitive to the highest advantage ratio, which here is infinity. Among the other indices considered, only Gr and L yield 50 percent for case C.

Conclusions

What have we added to Monroe's (1994) and Pennisi's (1998) earlier studies of disproportionality and malapportionment? First, we have updated and complemented the zoo of indices proposed and used by various researchers, pointing out parallel developments in measuring volatility and possible extension toward systematic measurement of vote splitting. Second, we have composed two sets of desirable criteria (theoretical and practical), rating each of the 19 indices explicitly on those 12 criteria. The most widely used indices, Gallagher's Gh and Loosemore-Hanby's D , satisfy more criteria than any other, including D_p , proposed by Monroe (1994).

Third, we have tried to specify a constellation that would be intuitively halfway from perfect concordance to utmost discordance; once more, Gh and D are among the few indices that yield 50 percent for such a constellation. We have located a dilemma when Dalton's principle of transfers is applied to parties rather than individuals: which party is the richer one when one has more extra seats while another is overpaid by a larger ratio? If the number of excess seats is taken as the preferred criterion, Gh becomes slightly preferable to D , in view of the latter's ambiguity when data sources lump several small parties. Finally, we tried to devise an even better index than Gh or D but came up short.

The reasons that have caused the recent shift from D toward Gh in the electoral studies literature boil down to sensitivity to party system concentration and interpretation of the paradox of transfers, which also may apply to volatility.¹³ The additional concerns mapped in this article should impact on studies of volatility once political party scholars recognize the analogies between the two types of measures we have pointed to above.¹⁴ We see no reason why Gh and D should not be the preferred indices for all four issues considered – deviation from PR, volatility of votes or seats, malapportionment and ticket splitting.¹⁵ Favoring some other index for any one of these topics would need justification to counterbalance the shortcomings listed here.

Notes

We are indebted to Cheryl Larsson and Clover Behrend for secretarial assistance, and to Kim DeFronzo for calling the helpful citation, White (1986), to our attention.

- 1 We are not aware of attempts to work out explicit discrepancy indices for vote splitting, beyond statements like 38 percent of Russian voters casting a split ticket vote in 1995, while the rate has been around 25 percent in the USA (McAllister and White, 2000). The issue becomes even thornier when three-way vote splitting is possible, say, among a lower chamber seat, an upper chamber seat and presidency.
- 2 When measuring the degree of discordance between two sets, perfect concordance is an obvious conceptual *extreme case*. In some cases, it may also be seen as a *normative standard*, but opinions may differ. Although most people might agree that excessive disproportionality, volatility, malapportionment and ticket splitting is undesirable, they may not always consider perfect concordance as desirable either. Indeed, many electoral rules work against excessive proportionality either explicitly (legal thresholds) or implicitly (low district magnitudes).
- 3 We consider only indices for deviation from PR that have been used or proposed in recent decades. Some of these have also been used for volatility, but we are not aware of any volatility measures that lack a disproportionality counterpart. Disproportionality is different from (though related to) inequality, where the base line is equality of shares. Inequality in turn differs completely from concentration (Taagepera, 1979). (For example, at the same zero level of inequality 50–50 is more concentrated than 25–25–25–25.) Deviation from PR is also related to malapportionment, but with some subtle differences. We do not consider indices such as Coulter's (1980) measure (which is close to D_p), or the index Fry and McLean (1991) adapted from a poverty measure, that have not been used to measure volatility or deviation from PR. These indices are, however, discussed by Monroe (1994). Monroe (1994), White (1986), Taagepera and Grofman (1981) and Taagepera (1979) contain important discussions of broader aspects of index creation, including the measurement of inequality, malapportionment and segregation.
- 4 Gh is our own notation for Gallagher's (1991) index.
- 5 Bernard Grofman has, in unpublished work, proposed that we use a closely related idea, the normalized sum of squares $\sum (y_i - x_i)^2 / \sum (Y_i - \bar{Y})^2$. This measure is a pseudo- r^2 . It shows the fit of the x values to the line $y = x$. This measure, however, can fall outside the (0, 1) range if disproportionality is extreme. If we take the square root of this measure, which is a pseudo- r value, we can see the direct parallel to Gh .
- 6 Note that the value of a depends upon whether we are regressing x on y or regressing y on x .
- 7 To permit PWI to be used with both percentages and shares we may apply a quasi-dimensional correction: $PWI' = PWI^{.5} = [\sum x_i / y_i - x_i]^{.5}$. Because the properties of PWI' are identical to those of PWI except that it satisfies the criterion of permitting calculations based on either percentages or shares, and it introduces additional complexity, we will not bother to discuss PWI' further.
- 8 We are concerned here with perfect concordance as a conceptual boundary case, regardless of what normative implications it may or may not offer.
- 9 For volatility and vote splitting, too, the largest difference seems to matter. However, for malapportionment one might hesitate. In particular, reducing a small component's share from one seat to zero might be considered more grave than a one-seat shift among large components.
- 10 For comparisons of D and Gh with respect to this 'lumping of residuals' criterion, see Appendix of Taagepera (1997).

- 11 For comparison of the forms of D and Gh in a quite different context, see Westholm (1997).
- 12 If the x - y symmetry criterion were considered unimportant, then *chi-square* and *T&S* would also become competitive.
- 13 As noted, the basic intuition is that a large change in votes (or seats) for one party should count more than small and possibly random shifts among several parties giving rise to the same cumulative discrepancy, even when the latter involve larger ratios between successive vote (or seat) shares.
- 14 We would also observe that, in addition to studying (average) election-to-election volatility, students of party volatility should also compare more distant elections. Sometimes election-to-election differences are picking up random aspects of electoral change and might be blind, say, to patterns of oscillation in power between two major parties (or groupings of parties) or to long-run patterns of realignment that are too subtle to catch over a short time frame. Empirical analyses relating to this point have begun by the present authors. Another measurement issue has to do with the scale on which analyses are performed. For example, the average disproportionality or volatility measured on the district level will not be the same as the total disproportionality or volatility observed in terms of national seat and vote shares (cf. Grofman et al., 1997).
- 15 A referee has suggested that volatility is different because a measure of volatility should be based on the total electorate rather than on the votes cast. True, one aspect of volatility is that people sometimes vote and sometimes don't. But then by the same token, shouldn't this 'party of non-voters' (that wins zero seats) also be included in calculations of disproportionality? The choice of appropriate database is separate from the choice of an index.

References

- Bartolini, S. and P. Mair (1990) *Identity, Competition, and Electoral Availability: The Stabilisation of European Electorates 1885–1985*. New York: Cambridge University Press.
- Coppedge, M. (1996) 'Latin American Parties: Political Darwinism in the Lost Decade'. Paper presented at the Conference on Political Parties and Democracy, National Endowment for Democracy, Washington, D.C., November 18–19.
- Coulter, P. B. (1980) 'Measuring the Inequity of Urban Political Services', *Policy Studies Journal* 8: 683–98.
- Cox, G. W. and M. S. Shugart (1991) 'Comment on Gallagher's "Proportionality, Disproportionality and Electoral Systems"', *Electoral Studies* 10: 348–52.
- Duncan, O. D. and B. Duncan (1955) 'A Methodical Analysis of Segregation Indexes', *American Sociological Review* 20: 210–17.
- Fry, V. and I. McLean (1991) 'A Note on Rose's Proportionality Index', *Electoral Studies* 10: 52–9.
- Galeotti, G. (1994) 'On Proportional Non-Representation', *Public Choice* 80: 359–70.
- Gallagher, M. (1991) 'Proportionality, Disproportionality and Electoral Systems', *Electoral Studies* 10: 33–51.
- Grilli di Cortona, P., C. Manzi, A. Pennisi, F. Ricca and B. Simeone (1999) *Evaluation and Optimization of Electoral Systems*. Philadelphia: SIAM.

- Grofman, B. (mid-1980s) Private communication, cited in Lijphart (1994: 61).
- Grofman, B., W. Koetzle and T. Brunell (1997) 'An Integrated Perspective on the Three Potential Sources of Partisan Bias: Malapportionment, Turnout Differences, and the Geographic Distribution of Party Vote Shares', *Electoral Studies* 16: 457–70.
- Laakso, M. and R. Taagepera (1979) '“Effective” Number of Parties: A Measure with Application to West Europe', *Comparative Political Studies* 12: 3–27.
- Li, J.-S. (1995) 'Analyse critique des indices de disproportionnalité électorale et de stabilité. Application aux cas de l'Allemagne et du Japon', *Revue internationale de politique comparée* 2: 369–88.
- Lijphart, A. (1984) *Democracies*. New Haven, CT: Yale University Press.
- Lijphart, A. (1994) *Electoral Systems and Party Systems*. New York and Oxford: Oxford University Press.
- Loosemore, J. and V. Hanby (1971) 'The Theoretical Limits of Maximum Distortion: Some Analytical Expressions for Electoral Systems', *British Journal of Political Science* 1: 467–77.
- McAllister, I. and S. White (2000) 'Split Ticket Voting in the 1995 Russian Duma Elections', *Electoral Studies* 19: 563–7.
- Monroe, B. L. (1994) 'Disproportionality and Malapportionment: Measuring Electoral Inequity', *Electoral Studies* 13: 132–49.
- Mudambi, R. (1997) 'A Complete Information Index for Measuring the Proportionality of Electoral Systems', *Applied Economics Letters* 4: 101–4.
- Nagel, S. S. (1984) *Public Policy: Goals, Measurement, and Methods*. New York: St. Martin's Press.
- Pennisi, A. (1998) 'Disproportionality Indexes and Robustness of Proportional Allocation Methods', *Electoral Studies* 17: 3–19.
- Przeworski, A. (1975) 'Institutionalization of Voting Patterns, or Is Mobilization a Source of Decay?' *American Political Science Review* 69: 49–67.
- Rae, D. W. (1967) *The Political Consequences of Electoral Laws*. New Haven, CT: Yale University Press.
- Taagepera, R. (1979) 'Inequality, Concentration, Imbalance', *Political Methodology* 6: 275–91.
- Taagepera, R. (1997) 'Effective Number of Parties for Incomplete Data', *Electoral Studies* 16: 145–51.
- Taagepera, R. and B. Grofman (1981) 'Effective Size and Number of Components', *Sociological Methods and Research* 10: 63–81.
- Taagepera, R. and M. S. Shugart (1989) *Seats and Votes: The Effects and Determinants of Electoral Systems*. New Haven, CT: Yale University Press.
- Westholm, A. (1997) 'Distance versus Direction: the Illusory Defeat of the Proximity Theory of Electoral Choice', *American Political Science Review* 91: 865–83.
- White, M. J. (1986) 'Segregation and Diversity Measures in Population Distribution', *Population Index* 52: 193–221.

REIN TAAGEPERA is Professor Emeritus at University of California, Irvine, where he has taught since 1970, and Tartu University, Estonia, where he has taught since 1992. His PhD was in physics, and his main research interest is in electoral systems. He is senior author of *Seats and Votes* (New Haven, CT: Yale University Press, 1989),

received 23 percent of the votes in the Estonian presidential elections of 1992, and was the founding chair (2001–2002) of a new Res Publica party that tied for first place in the Estonian parliamentary elections of 2003 and supplied the core of a coalition cabinet.

ADDRESS: School of Social Sciences, University of California, Irvine, 3151 Social Science Plaza, Irvine, CA 92697, USA. [email: rtaagepe@uci.edu]

BERNARD NORMAN GROFMAN received his BS in Mathematics at the University of Chicago in 1966 and his PhD in Political Science at the University of Chicago in 1972. He has been Professor of Political Science at the University of California, Irvine, since 1980. His past research has dealt with mathematical models of group decision-making, legislative representation, electoral rules and redistricting. Currently, he is working on comparative politics and political economy, with an emphasis on viewing the United States in comparative perspective. He is co-author of 2 books, both published by Cambridge University Press, and co-editor of 15 other books; he has published over 200 research articles and book chapters, including work in the *American Political Science Review*, the *American Journal of Political Science*, the *Journal of Politics*, the *British Journal of Political Science*, *Electoral Studies*, *Social Choice and Welfare*, and *Public Choice*. Professor Grofman is a past President of the Public Choice Society. In 2001 he became a member of the American Academy of Arts and Sciences.

ADDRESS: School of Social Sciences, University of California, Irvine, 3151 Social Science Plaza, Irvine CA 92697–5100, USA. [email: bgrofman@uci.edu]

Paper submitted 6 March 2002; accepted for publication 10 February 2003.