Title
Reformulating the Cube Law for Proportional Representation Elections

Journal
American Political Science Review, 80(2)

ISSN
0003-0554

Author
Taagepera, Rein

Publication Date
1986-06-01

DOI
10.2307/1958270

Peer reviewed

REFORMULATING THE CUBE LAW FOR PROPORTIONAL REPRESENTATION ELECTIONS

REIN TAAGEPERA
University of California, Irvine

The cube law *was proposed around 1910 to express the conversion of a party's vote shares into its seat share in two-party plurality elections with single-seat districts. This article develops predictive seat-vote equations for a much wider range of elections, including those involving many parties, single- and multi-seat districts, and diverse seat allocation rules such as plurality and list proportional representation (PR). Without any statistical curve fitting based on the seat and vote shares themselves, the basic features of the conversion are predicted using exogenous parameters: magnitude and number of districts, number of parties, and total size of the electorate and of the assembly. The link between the proposed equations and the original cube law is explicated. Using an existing data base, the fit of the predictive model is examined. On balance, this model accounts well for the conversion of votes to seats, and for the deviation from proportionality in PR systems.*

A fundamental issue for democracies is the extent to which elections reflect the popular will. One important aspect of that concern is the degree to which voter preferences for different parties and candidates are reflected in electoral outcomes. The way in which this question is commonly studied is by looking at the relationship between the aggregate vote share for candidates of a given party (or other grouping) and the aggregate seat share received by candidates of that party (Rae, 1971).

Some electoral systems expressly aim at proportional representation (PR), that is, seat shares equal to vote shares. However, the so-called PR systems differ widely among themselves in the degree to which they deviate from ideal PR, largely because different district magnitudes (i.e., the number of seats per district) are used, and different numbers of parties have evolved. Discussion of electoral reform in such countries often revolves not around the principle of PR, but around whether and how to make the system somewhat more or less proportional. However, such discussion has not been guided by any general quantitative rule predicting the specific degree of deviation from PR in the systems existing or proposed.

For plurality elections, it is well known that proportionality between vote shares and seat shares cannot be expected. In this case, however, a votes-to-seats conversion rule has been proposed, the so-called cube law. Devising a comparable expression for the PR elections would have considerable practical importance in guiding the discussions of electoral reform. A more general expression accounting for votes-to-seats conversion in both plurality and PR elections would also be of theoretical significance in unifying the field of electoral systems, and in casting new light on the functioning of each particular system through quantitative comparison with others.

The purpose of this article is to develop and test a set of predictive equations that use no curve fitting based on the seat and vote shares themselves. Starting out as a reformulation and expansion of the cube law, these new seat-vote equations express the basic features of conversion of vote shares into seat shares for a number of electoral rules, including plurality, "list PR" elections, and the Japanese single nontransferable vote.

The Cube Law of Elections

For two-party contests in single-seat ("single-member") districts using plurality, the *cube law of elections,* first formulated around 1910 (see Kendall and Stuart, 1950) was proposed as a predictor of the expected seat-vote relationship. The cube law expresses the empirical observation that cubing the ratio of votes (v_K/v_L) received by any two parties K and L was seen to approximate the ratio of assembly seats (s_K/s_L) they win:

$$s_K/s_L = (v_K/v_L)^3. \qquad (1)$$

What this means is that any advantage a party has in terms of votes is magnified at the seats level, and in a very specific way.

The general form taken by the cube law is unusual in political science (though quite usual in some other sciences), in that it consists of an equation connecting well-defined quantities, without any adjustable parameters subject to statistical curve fitting. Given its empirical nature, it might more properly be called the "cube rule," but the traditional designation will be kept in this article. The cube law is meant to apply only to electoral systems of the Anglo-Saxon type that use one-seat districts and the plurality ("first past the post") allocation rule. In particular, it is not expected to apply to PR systems in multi-seat districts. The latter try to achieve something close to $s_K/s_L = v_K/v_L$, thus replacing the power index 3 in equation (1) with 1.

Even some Anglo-Saxon elections deviate considerably from the cube law (Tufte, 1973). They involve partisan gerrymander, which biases the pattern in favor of one party, and bipartisan gerrymander, which profits incumbents in both major parties. Despite such occurrences the cube law might still be a useful simple base-line rule. It could supply a framework to which various correctives could gradually be added—e.g., some bias or gerrymander coefficients. (For a recent overview of measures of bias see Grofman, 1983.)

The predictive power of the cube law is more seriously threatened by observations that in many cases a power index other than 3 gives a better fit. Even though the cube law was initially formulated for Great Britain, some recent work suggests that a power index of 2.5 might fit British elections better (Laakso, 1979). Further analysis (Schrodt, 1980) underlines the difficulty of establishing the proper statistical format for testing such a nonlinear expression.

Previous Generalizations of the Cube Law

A number of ingenious theoretical explanations for the existence of the cube law have been offered. Most of them (see appendix 1) are so specific that they have not offered avenues for testable predictions beyond the Anglo-Saxon parliamentary elections. However, there is one previous line of approach which has yielded testable predictions for some other elections (such as the U.S. Electoral College), but always within the plurality framework. This approach will be reviewed here, because it will supply the starting point for further generalization to multiseat districts, to be presented in the next section.

The properties of the following equation were studied by Theil (1969):

$$s_K = v_K{}^n / \Sigma\, v_i{}^n, \qquad (2)$$

where s_K and v_K are one particular party's vote and seat shares, respectively, n is a constant, and the summation is over the vote shares of all parties. This equation expresses the seat share of one specific party, K, in terms of the vote shares of K and all other parties. When dividing equation (2) by the analogous equation for party L, the summation term cancels out, and one obtains

$$s_K/s_L = (v_K/v_L)^n. \qquad (3)$$

This form would include, as special cases, the cube law (for $n=3$) and perfect PR (for $n=1$). Equation (2), in turn, can be derived from equation (3), which means that the two forms are mathematically equivalent. In the case of the cube law, the summational form was used by Qualter (1968) for Canadian elections.

Theil (1969) showed that if any equation connecting s_K/s_L to v_K/v_L exists at all, it has to be of the general form given above. The reason is that the sum of the seat shares of all parties must amount to unity, or 100%. Summing the seat shares s_i for all parties, using equation (2), does yield such a result. This could not be the case for any other functional relation between s_K/s_L and v_K/v_L, when the system has three or more parties, according to Theil's mathematical proof. What this means is that equation (3) is the only possible expression connecting seat and vote ratios that does not produce inconsistencies.

At this stage the index n could have any value, and the question remains why it should be equal to 3 in the case of national assembly elections. It has been proposed (Taagepera, 1973) that n depends on the logarithms of the total number of votes (V) and the total number of districts (D):

$$n = \log V / \log D, \qquad (4)$$

provided that all seats in the same district are allocated to the party with plurality of votes.

The broad underlying reason is as follows. When one considers the ideal extreme cases $D=V$ and $D=1$, it becomes apparent that the numerical value of n should be affected by the relative number of votes and districts. If districts are made so numerous that every voter is a separate district ($D=V$), then perfect PR should prevail by definition, so that $n=1$. If, on the other hand, the number of districts is reduced to one (as in direct presidential elections), then only one party can win a seat, and the resulting seat ratio 0:1 emerges from equation (3) only if n tends to infinity. Assuming quasi continuity in the intervening range, one would expect that n gradually increases from 1 to infinity as D gradually decreases from V to 1.

Equation (4) expresses this general trend and also satisfies the boundary conditions at $D=1$ and $D=V$. The theoretical/mathematical justification for the logarithmic format is reviewed in appendix 2. The striking aspect of equation (4), however, is that it does agree with actual elections, and not only in the cases where the cube law might apply, but also in U.S. Electoral College elections and certain trade union elections data.

For most national assembly elections, the ratio log V/log S, where S is the number of assembly seats, happens to be close to 3, regardless of electoral rules used (see Table 1 for some examples, disregarding for the moment the last three columns, M, N, and n). In other words, a cube root relationship tends to exist worldwide between assembly size and the voting population, and a rational model for such a relation has been proposed by

Table 1. Characteristics of Electoral Systems Selected for Model Testing

Country, Period and Number of Elections	Votes (V, millions)	Seats ($S=DM$)	$\frac{\log V}{\log S}$	District Magnitude (M)	Effective Parties (N)	n
U.S. House 1950–70 (11)	54.9	436	2.93	1	2.0	2.93
Canada 1963–74 (5)	8.3	264	2.86	1	3.1	2.86
Japan 1963–76 (5)	47.5	486	2.86	4	3.3	1.30
Austria 1945–70 (8)	4.4	165	3.00	6	2.5	1.20
Switzerland 1947–75 (8)	1.0	200	2.61	8	5.4	1.13
Italy 1958–76 (5)	30.2	630	2.67	19	3.9	1.05
Finland 1962–75 (5)	2.5	200	2.78	15	5.8	1.07
Netherlands 1963–77 (5)	6.3	150	3.12	150	5.4	1.01

Sources: Median V, S, and N calculated from data in Nohlen (1978, pp. 384–415), except for the U.S. (Vick, 1979, based on Cox, 1972). Average M is calculated or estimated from information in various parts of Nohlen (1978). Equation (5) is used to calculate n. The number D of districts resulting from $S=DM$ may differ slightly from the actual number when M is not the same for all districts.

Taagepera (1972). In the case of single-seat districts (where $D=S$) this means that the value of n resulting from equation (4) is close to 3. The cube law now appears as the result of combining the "cube root law of assembly sizes" with the equations (3) and (4).

While the cube law is thus explained in terms of more basic rules and thus becomes more credible, it also is subject to a new unexpected modification: instead of being exactly 3 in all cases, the exponent is now expected to vary somewhat from country to country (and also over time within the same country). For New Zealand, equation (4) yields values around $n=3.2$, while for Britain it yields $n=2.6$, with intermediate values for Canada (2.8) and the U.S. (2.9).

Besides casting such new light on the cube law, equations (3) and (4) also lead to predictions regarding the U.S. electoral college, where $D=51$ (and less in earlier times), given that all delegate seats of a state go to the same party. Instead of a power index close to 3, equation (4) now yields an average of about 5, and this is indeed close to the observed average swing ratio, which is clearly higher than predicted by the cube law (Taagepera, 1973). Certain trade union elections studied by Coleman (1964, p. 347) also fit the pattern, but there the calculated (and observed) n is much less than 3: namely, around 1.5 (Taagepera, 1973). As observed earlier, equation (4) also applies to direct presidential elections, where the entire country becomes a single one-seat district.

In sum, the combination of equations (3) and (4) can express and predict the broad features of Anglo-Saxon assembly elections, of U.S. Electoral College elections, and of direct presidential elections. The fit to any particular election results may be far from the best statistical fit, but the latter lack any predictive ability regarding other countries.

Apart from direct testing or theoretical proof, a model's credibility increases if it leads to unanticipated consequences that can be observed. The application to U.S. Electoral College and direct presidential elections is of that type, since both fall clearly outside the scope of the original cube law. The usefulness of a model is also enhanced if it ties together several previously isolated phenomena. Combining the aforementioned "cube root rule of assembly sizes" (as proposed by Taagepera, 1972) and the seat-vote relationships in equations (3) and (4) to yield the cube law of elections is such an extension.

Like the cube law, the generalized seat-vote relationship thus developed is meant to apply only to systems using the plurality allocation rule. We will now proceed to previously uncharted ground, and extend the seat-vote equations to multi-seat districts with PR allocation rules. It is sometimes thought that such an extension is unnecessary, because "PR is PR." However, actual PR systems deviate from the ideal PR, and they do so to varying degrees. It will be seen that our extended model is able to account for the basic differences between them.

Seat-Vote Equations for Multi-Seat Districts with Multiple Parties

Further extension of the model of seat-vote relations to include PR elections in multi-seat districts first requires the introduction of district magnitude, M, i.e., the number of seats allocated in the same electoral district using some PR or other non-plurality formula. On grounds that will be discussed shortly, it is proposed that magnitude affects the power exponent n and requires the following modification of equation (4):

$$n = (\log V/\log DM)^{1/M}. \tag{5}$$

In the case of single-seat districts ($M=1$) this equation reduces to the earlier one, as it should. In the case of very large districts (such as Netherlands, where the whole country is a single electoral district of $M=150$) n tends toward unity, reflecting near-perfect PR. For low-magnitude, multi-seat districts (such as Japan, where $M=4$), the small party penalty inherent in the cube law is reduced to a considerable extent, but the system still is a long way from perfect PR.

The way M is introduced is chosen so as to satisfy the boundary conditions for the extreme cases of $M=1$ (cube law) and M tending toward infinity (perfect PR), but beyond this it is purely empirical. It is justified to the extent that it correctly predicts the broad patterns of the seat-vote relationship for a wide range of district magnitudes. The equation would be expected to apply best to systems with a simple electoral setup: having all districts be the same magnitude and using a simple allocation rule such as d'Hondt or Quota. The complexities actual electoral systems can exhibit are almost limitless (for an overview, see e.g. Grofman, 1975, or Lijphart, 1984a). Thresholds, nationwide adjustment seats, use of second preferences (through ranking on the ballot or through a second round of voting), high variance in district magnitude, and other such features are not taken into account by equation (5). Since such features should be expected to reduce the predictive power of our model, if it is nonetheless able to account well for the basic features of votes-to-seats conversion, our confidence in the model's usefulness would be greatly enhanced.

Used in conjunction with equation (2), equation (5) would enable us to calculate the seat shares s_K for any given value of v_K, except for one remaining problem: the outcome does not depend only on the given party's share of votes, but also on how the remaining votes are distributed among the other parties, through the summation in the denominator of equation (2). In plurality elections, in particular, votes distribution among other parties matters greatly. A vote share of 40% can spell disaster if the party is facing a single opponent, or landslide victory if it is facing two opponents of equal size (as Thatcher did in 1983).

The main issue here is the number of serious opponents one faces. To operationalize "serious" we will use the "effective number of parties," as defined by Laakso and Taagepera (1979), and recently used by Lijphart (1984b) to establish major conclusions regarding the functioning of democracies:

$$N = 1/\Sigma v_i^2, \quad (6)$$

where the summation is over all parties in the given elections.[1] While N changes from election to election, the changes are usually observed to be minor, so that the average N over a long period can be used (Laakso and Taagepera, 1979; Lijphart, 1984b). Using such an average value, we shall approximate the denominator in equation (2) by assuming that every party K faces $N-1$ other parties of equal size in terms of votes. This gives us, for the denominator of equation (2),

$$\Sigma v_i^n = v_K^n + (N-1)^{1-n}(1-v_K)^n. \quad (7)$$

This seemingly complex expression actually represents a simplification, in the sense that instead of having to know the vote shares of all the parties, one only needs to know N and the vote share v_K of the single real or hypothetical party under consideration. In sum, the seat-vote equations for single-seat plurality and multi-seat PR elections will be tested in the form

$$s_K = v_K^n/[v_K^n + (N-1)^{1-n}(1-v_K)^n], \quad (8)$$

where n is given by

$$n = (\log V/\log DM)^{1/M}. \quad (5)$$

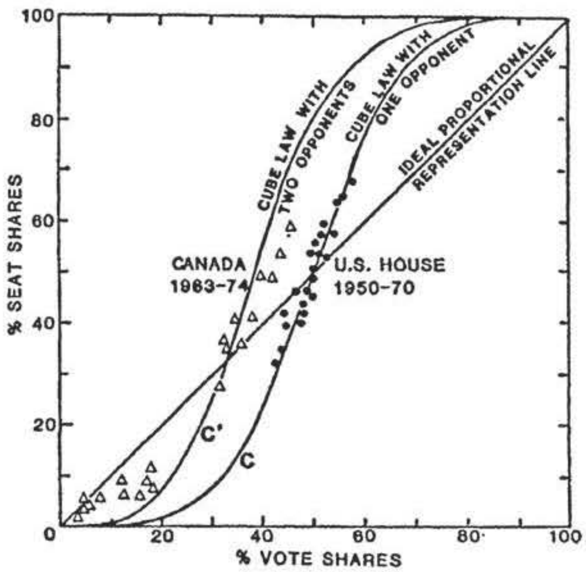
We can now calculate the seat share for any given vote share, provided that N and n are known. The latter in turn depends on the average district magnitude, the number of districts, and the number of votes in the nation. Thus we are predicting the shape of the relationship between seats and votes in terms of one factor (V) independent of the electoral rules, two factors stipulated in the electoral rules (M and D), and a factor (N) which summarizes the long-term votes distribution. If those few factors exogenous to the election results of the given year enable us to account for the general pattern of votes-to-seats conversion in a wide range of multi-seat PR and single-seat plurality systems, we would be doing well, indeed.

Graphical Analysis of the Seats-Votes Model

The commonsense way to graph the seat-vote relation for various parties is simply to plot the seat share of party K (s_K) versus its share of votes (v_K), as shown in Figure 1 for Canada and the U.S. However, this way of graphing results in an overcrowding of points at the lower left and a lack of sensitivity regarding even major deviations from PR. Curves C and C′ express the theoretical outcomes calculated from equation (8) for the U.S. House (N = 2.0) and for Canada (N = 3.1), respectively. Given that n is close to 3 in both cases (see Table 1), these curves essentially express the outcome of the traditional cube law for parties facing a single opponent or two opponents of equal size, respectively. Although the cube law pattern is a standard example of a gross deviation from PR, the curves C and C′ do not visually seem to deviate that much from the line of perfect PR. Most actual data points for elections world-wide are squeezed in between the curves shown and the ideal PR line. Even stark deviations from PR would be deemphasized both in visual inspection and in statistical testing. Thus the commonsense way to display seats-votes data is not very good.

The crowding and loss of discrimination are avoided if, instead of s_K, one plots the "advantage ratio" $A_K = s_K/v_K$ versus v_K, a format used earlier by Taagepera and Laakso (1980). The outcomes for Canada and the U.S. are shown in Figure 2. The perfect PR and cube law curves now become very distinct, and considerable detail emerges for the small Canadian parties. In the first plot (s vs. v) they all seem to fall close to the PR line. In the second plot (A vs. v) some visibly fall

Figure 1. A Non-Optimal Way to Plot Electoral Data: Seat Shares vs. Vote Shares



very much short of PR, while in one case there is considerable overrepresentation. Note that such "proportionality profiles" have one point for every party in every election.

To demonstrate how widely disparate proportionality profiles can be engendered by equation (8), depending on district magnitude, a sampling of theoretical curves is shown in Figure 3.[2] These curves show that, in comparison with PR rules, the plurality rule ($M=1$) is expected to give a greater relative advantage to large parties, as has long been observed empirically. They further show that within the PR systems the large party bonus should decrease as M increases, a point previously noted by Sartori (1968) among others.

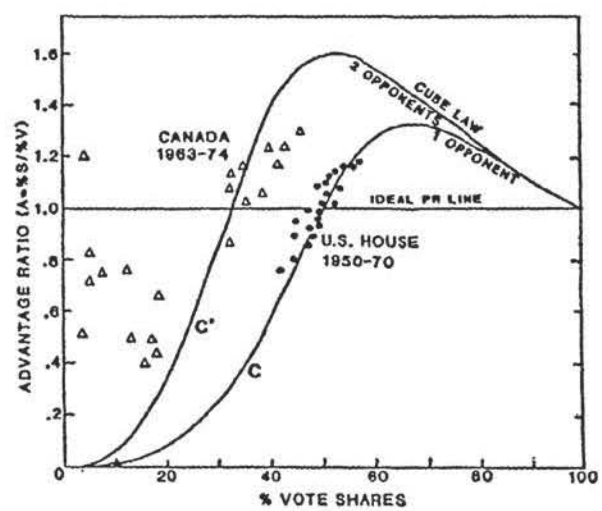
Graphical Testing of the Seat-Votes Model

While one might eventually want to test the fit of the proposed equations for all available electoral data, all that is really needed is to look at a data base that adequately covers the entire observed range of values of both district magnitude ($M=1$ to 150) and the number of parties ($N=2$ to 6). To reduce the possibility of picking only the cases which best agree with the model, the selection of countries and periods was based on two previously existing works: Lijphart's *Democracies* (1984b) and Nohlen's *Wahlsysteme der Welt* (1978).

Case Selection

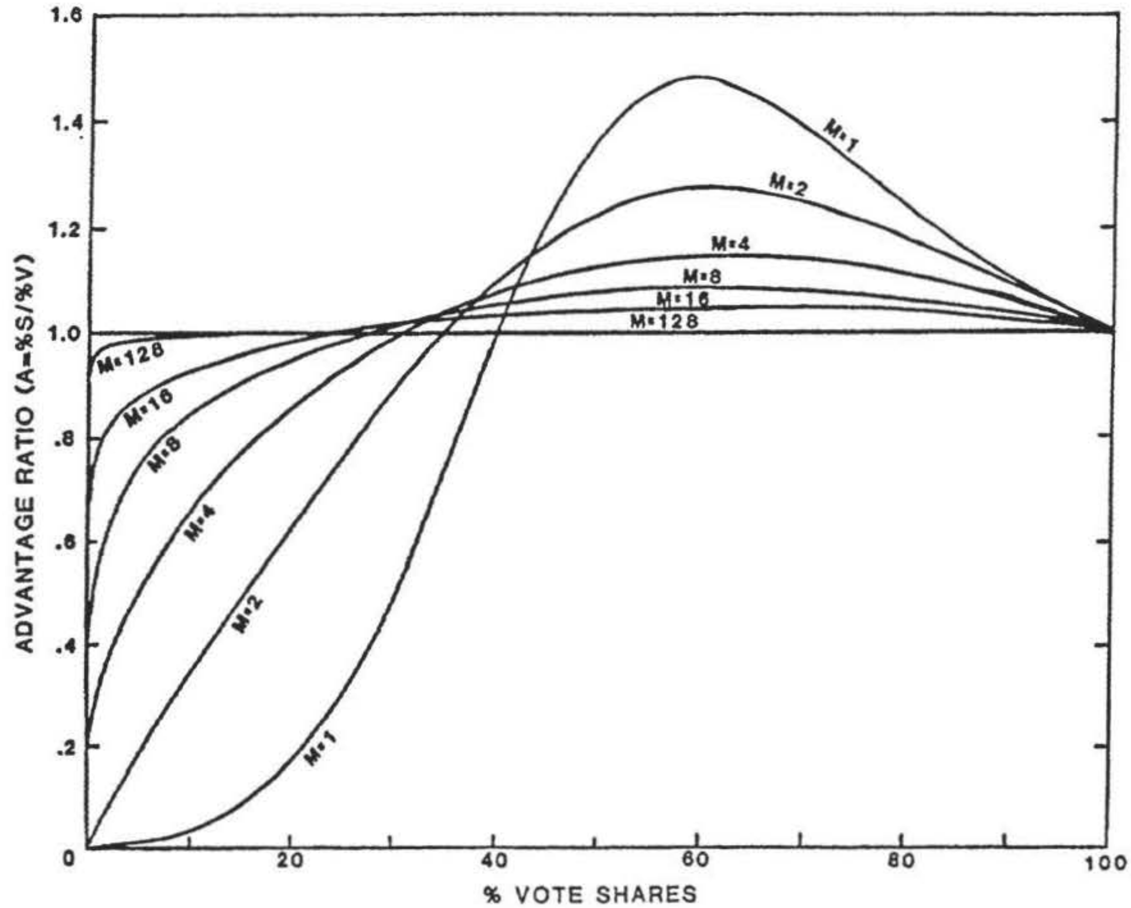
All electoral systems among Lijphart's (1984b) list of 21 stable democracies were considered, discarding those which involve extensive nationwide or regional adjustment seats, rather high thresholds, a second round, use of second preferences, or mixing of several allocation rules. Among the remaining 14 cases, those with the lowest and the highest effective number of parties were chosen for different ranges of district magnitude.[3] For the 8 cases thus selected (see Table 1), the election data to be used were chosen on the basis of elections tabulated in the

Figure 2. Proportionality Profiles: Advantage Ratios vs. Vote Shares



Note: See Table 1 for data sources and parameter values of the predicted curves shown in Figures 1 and 2.

Figure 3. Predicted Proportionality Profiles at Various District Magnitudes (*M*)



Note: These profiles have been calculated using equation (8), with $n=3^{1/2M}$ and $N=1.25(2+\log m)$.

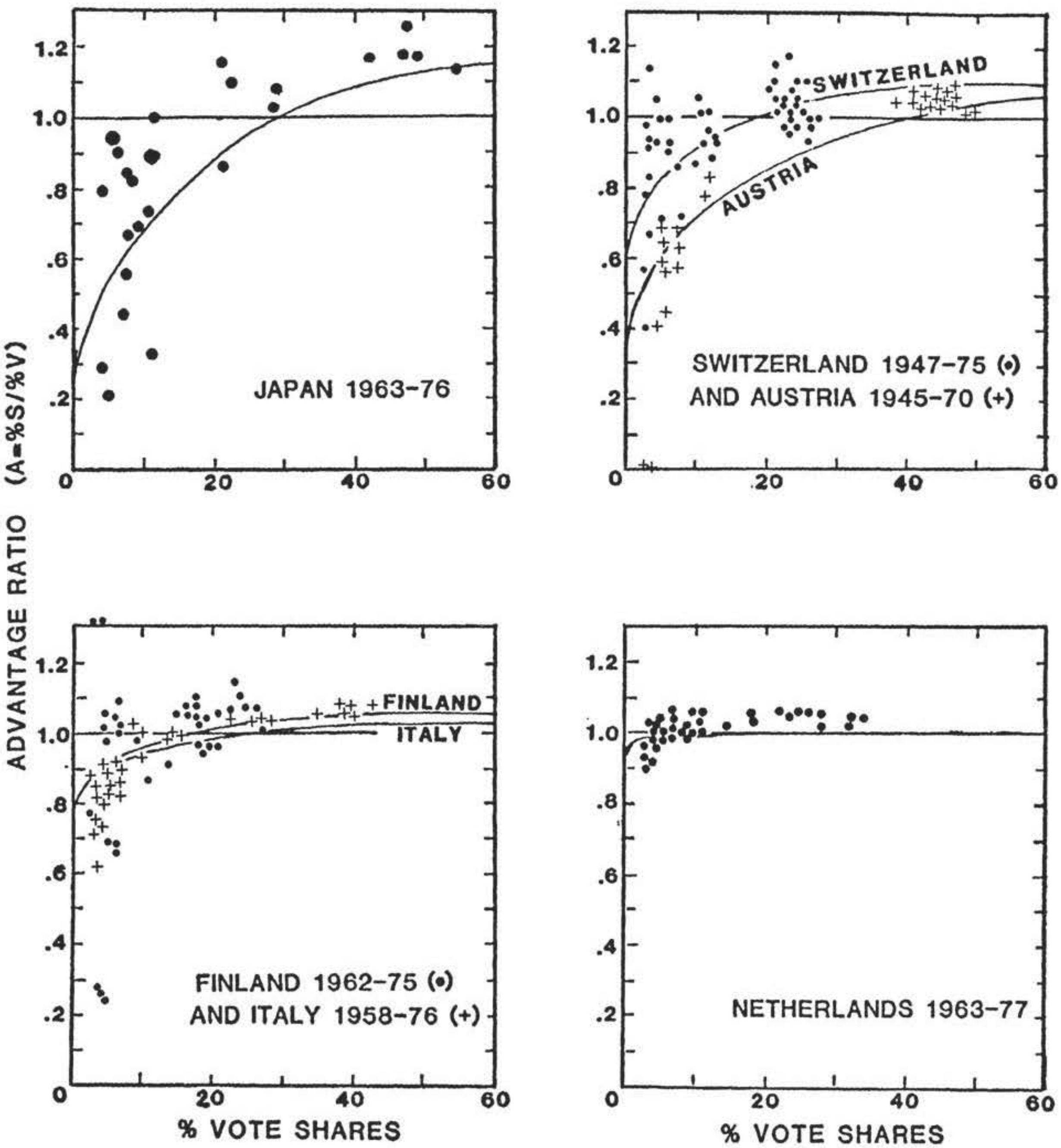
appendix of Nohlen (1978).[4] Average magnitudes were calculated or estimated from information in Nohlen, and rounded off to integers. The average *n* and *N* were calculated using equations (5) and (6). The expected seat shares at various vote shares were calculated using equation (8), and the resulting theoretical proportionality profiles were graphed. Figures 2 and 4 show the actual results for all parties in all the elections considered, plotted superimposed to the theoretical curves.[5]

In assessing the fit of the model to the data from the eight countries, it should be kept in mind that this is *not* a curve fitting with four coefficients, the values of which can be adjusted at will to achieve the best fit. The values of *M, N, V,* and *D* are all predetermined by the political system. There are no adjustable fudge factors. This also means that there is no need for an "after-the-fact test of the fitted model" with a different data set, because there was no fitted model to start with: all coefficients are exogenous to the data on seats and votes.

Country Results

United States. Among the single-seat systems, the U.S. House offers the lowest number of parties ($N=2.0$). The fit of the U.S. data to the calculated curve in Figure

Figure 4. Predicted and Actual Proportionality Profiles for Elections with Multi-Seat Districts



Note: See Table 1 for data sources and parameter values used.

2 (which closely agrees with the classical cube law) is visibly close to the best possible fit.

Canada. Among the single-seat systems, Canada has the largest effective number of parties ($N=3.1$). While in the less sensitive s vs. v plot in Figure 1 the Canadian data may look close to the calculated curve, the more sensitive proportionality profile (A vs. v plot) in Figure 2 reveals considerable deviation, especially for the smaller parties. Like the classical cube law, the present refined form correctly predicts very few or no seats for small nationwide parties, such as the British Liberals, but it cannot account for regionally based parties such as the Scottish Nationalists, for whose success the nationwide vote share is not relevant. In such cases A actually tends to decrease with increasing vote shares, up to about 15% of the votes, reflecting the unsuccessful attempts of some regionally based parties to go nationwide. Canada offers one of the most marked cases of such regional variety.

Japan. Figure 4 shows various patterns for multi-seat districts. Japan is the only case available with a low M and a simple allocation rule, and even there the rule is rather unusual: the single nontransferable vote. The calculated curve tends to fall below the actual points, but it correctly predicts the general pattern: (1) The penalty for nonregional small parties is less than in the case of single-seat districts, but is still considerable. (2) The bonus for the large parties is less than in the case of single-seat districts, but it is still higher than in any other multi-seat system (as we will see). (3) The shift from penalty to bonus occurs at lower vote share levels (less than 30%) than is the case for single-seat districts (32–45%). As in the case of nationwide Canadian parties, the general shape of the proportionality of profile is well predicted for Japan, too.

Austria and Switzerland. The next panel in Figure 4 shows the calculated curves and election data for two countries with medium district magnitudes in which the number of parties differs widely: Austria ($N=2.5$) and Switzerland ($N=5.4$).[6] At a given percentage of votes, the calculated curves predict higher values of A in the case of Switzerland. This is confirmed by electoral results. The imperfect detailed fit should not detract from this main fact: as N changes (at essentially constant M) the direction and the approximate amount of change in A are correctly predicted. Furthermore, the maximum advantage ratio for large parties is predicted to be less than that for Japan, because of larger district magnitude, and this is also borne out. The small-party penalty is correctly predicted to be about equal in Austria and Japan, but less heavy in Switzerland. The shift from small-party penalty to large-party bonus occurs in Switzerland at a lower percentage of votes (around 20%) than in Japan, as predicted by the curves; in Austria's case the lack of parties with vote shares between 20% and 35% leaves this question open.

Finland and Italy. These countries have very large district magnitudes (short of a single countrywide district).[7] Although these countries differ considerably in the number of their parties, the calculated curves predict that at such a large magnitude N makes hardly any difference. If anything, the Italian data points would be expected to lie slightly lower than the Finnish, and on the average this is the case. Compared with Switzerland, the calculated curves for Finland and Italy predict a slight further reduction in small-party penalty, but this small difference cannot be ascertained with the actual data, given their wide random scatter. The prediction of a slightly lower large-party bonus, compared to that of Austria and Switzerland, is neither confirmed nor disconfirmed, given the scatter in data points. The predicted curves are by now so close to the ideal PR line that distinctions cannot be made regarding the break-even point. Once again the predicted curves tend to lie slightly too low for the larger parties.

The Netherlands. For the huge magnitude ($M=150$) of the Netherlands' countrywide district, the calculated curve is practically indistinguishable from the ideal PR line. The actual data confirm this linearity, but the best-fitting horizontal line lies at $A = 1.03$ rather than $A = 1.00$. (This is compensated by heavy penalty on parties with less than 2.5% of the votes, which are not shown in the plot.) The prediction of an even lesser large-party bonus and small-party penalty than in the case of Finland and Italy is borne out, but the difference is small, as expected from the model, and could be swamped by random noise.

Statistical Considerations

The degree of agreement of the predicted curves with the actual data readily meets the coarse but hard-to-fool test called eye-balling. This is true, in particular, regarding predictions about what would happen when district magnitude is increased, or when, at the same M, the number of parties is increased. A more formal statistical analysis will be given next.

For each country, the data were grouped by vote share brackets of 10%—i.e., up to 9.9%, from 10.0% to 19.9%, etc. Within each bracket the median vote share and the median advantage ratio A were determined. The theoretical advantage ratio (designated in this section as A') was also calculated on the basis of the median vote share, using equation (8) and the appropriate country parameters (M, D, V, and N). If the model fitted, one would expect not only a high correlation between the calculated A' and the actual value of A, but also a best-fit line close to $A = 1.0\,A' + 0.0$. This is so, indeed, especially in the case of multi-seat districts. For the 24 number pairs (A and A') generated by the 6 multi-seat systems, the predicted A' ranges from 0.6 to 1.1; the best-fit line is very close to $A=A'$, namely $A = .92A'+.13$, and $r^2=.87$.

The goodness of this fit is highlighted when one considers the best empirical predictions based on Rae's (1971, p. 89) linear fit of the percentage of seats to the percentage of votes for a large number of PR elections ($s = 1.07v - .84\%$).[8] This equation yields advantage ratio predictions which are not only rather widely scattered ($r^2 = .54$), but also congregate along a best-fit line with a slope deviating from unity: $A = 1.65A'-.66$. Using the best linear fit of seats to votes in our actual testing sample for a kind of "post-diction" offers no improvement: $A = .55A'-.46$; $r^2=.44$. For the multi-seat districts, our model performs vastly better than the empirical approach.

When including the 6 number pairs generated by the two single-seat systems in our sample, the goodness of fit is reduced by the aforementioned small, regionally based Canadian parties. The range of predicted values A' widens (from 0.02 to 1.4), the best-fit line for our model becomes $A=.54A'+.48$, and r^2 is .75. Small but nationwide parties in single-seat districts (e.g., in the U.K. or New Zealand) would follow the model, but the U.S. has no third parties with more than 2.5% of votes. This leaves Canada as the only plurality system in our sample which has small parties, putting the model to the severest possible test. Nonetheless, our model still does appreciably better than empirical predictions based on Rae's (1971, p. 70) fit of a large number of elections with single- and multi-seat districts ($s = 1.13v - 2.38\%$). The latter yields a correlation between expected and actual values of A of ony $r^2=.32$. The best fit of the percentage of seats to the percentage of votes for our present sample ($s = 1.15v - 2.45\%$) is very close to Rae's, and its "postdicted" advantage ratio values have the same limited correlation ($r^2 = .32$) with the actual A.

The elections held after the compilation of the Nohlen (1978) data are too few to offer a sufficient data base for a separate testing of the model. If, however, 13 later elections in the same countries are included with the previous data base of 52 elections, the outcome remains essentially the same, with $r^2 = .70$ for predictions based on our model and applied to all countries, including Canada.[9]

Conclusions

The qualitative predictions of the seats-votes model expressed in equation (8) are confirmed in the cases where large differences in proportionality patterns were expected. The deviations from proportionality predicted by our model agree with the actual ones much better than is the case for predictions based on any other approach (short of country-by-country curve fitting): $r^2 = .87$ for our model vs. .54 for the best alternative approach in the case of PR in multi-seat districts, and $r^2 = .75$ for our model vs. .32 when including plurality elections. Of course, better empirical fits could easily be found for each country separately, but such post-fact curve fitting would be something quite different from the approach used here, since it would be of no use in predicting outcomes in countries apart from the one for which it was established. The major achievement of the present approach is to enable one to make baseline predictions for almost any electoral system for which the number and magnitude of districts and the number of voters and parties are well definable and known.

The problem of expressing the relationship between seats and votes has of course not been fully solved. There are electoral rules that defy any systematics, such as the stipulation that the opposition parties to the PRI are to receive 25% of the seats in the Mexican national assembly, regardless of their percentage of votes. Other phenomena, such as vote size thresholds and mixing of several allocation rules, could in principle be taken into account by developing appropriate correction terms to the model presented here. The same applies to the effect of PR allocation rules (such as d'Hondt and Largest Remainder) at the same district magnitude and number of parties, although their effect is small compared to that of M and N.

How does the model presented agree with countries and periods other than the ones tested in this article? Given our method for selecting the sample countries, other systems with simple allocation rules should fit equally well. Visual inspection of previously published West European proportionality profiles (Taagepera and Laakso, 1980) confirms this expectation. The partly theory-based and partly empirical equations presented here clearly offer sufficient potential to warrant further testing with all electoral data available. They have predictive power, and by the same token are clearly "falsifiable"—an indispensable feature of scientific models and laws. For systems with complex electoral rules they offer a way to tell to which simpler system (if any) such a system is equivalent, as far as the votes-to-seats conversion is concerned.

In sum, with four parameters (M, N, V, and D), which are themselves exogenous to the elections data, we can account very well for seats-votes relationships in a very wide range of electoral and party systems. The form of interaction between these parameters is derived from theoretical considerations about their functional relationships. Given the exogenous nature of these parameters, the proposed seat-vote equations offer considerable interest for "electoral engineering," since the model can make specific predictions about how the votes-to-seats conversion would change upon, say, a small alteration in average district magnitude. But the main interest is intellectual and conceptual. Regarding the multi-seat districts, the small but significant deviations from PR can now for the first time be "understood" in a sense that includes quantitative prediction. For plurality in single-seat districts the known seats-votes relationships have been clarified by making them part of a broader picture. To an appreciable extent, single-seat and multi-seat electoral systems have up to now been considered distinct species to be studied separately. Unifying the field of electoral systems through equations which apply to both species certainly has theoretical significance.

Appendix 1. Previous Explanations of the Cube Law

A normal distribution of voters for a party over all districts would yield the cube law, provided that the appropriate value is used for variance (Kendall and Stuart, 1950). The general shape of the relationship thus receives an explanation, but the particular exponent value, 3, does not. The suitable value of variance emerges from a contagious binomial model based on a Markov chain approach, provided that people form reciprocal influence clusters of a suitable size (Coleman, 1964, pp. 343–53). The establishment of a tie between cluster size and the outcome of an election is of considerable interest in its own right, but it does not explain why people aggregate in clusters of precisely the right size to yield a cube relationship in national elections.

A Poisson distribution can be used instead of the normal one, and it can be justified in terms of a model of interaction between interparty negotiations and intraconstituency pressures (March, 1957); however, the underlying factors have not been measured. A LOGIT approach (Theil, 1970) also leads to the desired curve shape; the cube relationship emerges if one arbitrarily assumes that the standard deviation of the vote proportions across districts is around 1/8. A stochastic model by Quandt (1974) chiefly adds the possibility of bias in favor of one of the parties in a two-party system. The large number of parameters enables the model to account, in retrospect, for a large variety of shapes of the seats-votes relationship, but at the cost of making prediction unfeasible. The Markov scheme has been further elaborated upon by Gudgin and Taylor (1979, pp. 31–53) who may have been unaware of Coleman's work. The cluster sizes that would give rise to the cube law are not independent input data, but again are calculated retroactively from the cube law.

A quite different game theory approach to campaign resource strategy is followed by Sankoff and Mellos (1972); it leads to an exponent of 2 rather than 3. Additional assumptions regarding an unswayable hard core of partisan voters can lead to an exponent of 3—or any other value. Once more we have an interesting connection, but a hard-core percentage calculated from the cube law cannot pass for an explanation of it.

Broadly speaking, the aforementioned approaches tend to show the plausibility rather than the inevitability of the cube law. Some of them enable us to use the exponent 3 to infer what the prevalent cluster size or hard-core percentage would have to be, but they do not explain why the whole magic package comes about. The very number of distribution-oriented models offered (which include all those mentioned, except Sankoff and Mellos) indicates some unease regarding the preexisting models.

What would constitute an adequate explanation of the cube law? First of all, one would ideally want a model plus some independently known quantitites such that one could plug those pre-given numbers into the model from the one end and have equation (1) pop out from the other end. The variance or standard deviation of the distribution of votes over districts can be independently known (though in practice this is often difficult), and thus models using this basis would satisfy this first condition. As a further desideratum, one would prefer a model which applies to all parties, all vote ranges, and to any number or size of districts. All the aforementioned explanations use a system with two parties only. If they can be expanded to include nationwide third parties such as the British Liberals, this aspect has not been formulated. Models which apply to pure two-party systems are not without merit, but multi-party models including the two-party systems as a special case offer a wider explanation. The approach taken by Theil (1969) leads in that direction.

Appendix 2. Derivation of $n = \log V/\log D$[10]

In the main text of this article, reasoning through extreme border cases led to the conclusion that n is a function of the total number of voters (V) and of districts (D): $n = n(V,D)$. For the sake of achieving as much generality as possible, we would want the model to apply also in the case of multi-stage elections, such as the voters/electoral college/president sequence, although such elections occur rarely. We are here again reasoning by extreme cases which may or may not occur in practice. But if they were ever to occur, a general model of elections should be able to fit them without running into inconsistencies. If V voters elect D sub-representatives who in turn elect T final representatives, then one would expect equation (3) to apply to the stages V-D and D-T separately, and also to the entire process V-T. Consistency then requires that $n(V,T) = n(V,D)n(D,T)$. We can rewrite it, provided that $n(D,T)$ is not equal to zero, as

$$n(V,D) = n(V,T)/n(D,T).$$

Now take T to be a standard number of units, against which elections with any other number of units are compared. Then $n(V,T) = n(V,\text{constant}) = f(V)$ only, and similarly, $n(D,T) = f(D)$, so that

$$n(V,D) = f(V)/f(D).$$

The function f must be the same function for V and for D. Also, we must have $f(1)=0$, if we want to account for presidential elections.

In the absence of any information to the contrary, we have to assume that the degree of clustering of like-minded voters is the same on all levels (i.e., a magnified piece of a political adherence map would look like the original map). When V voters elect D representatives (with V/D voters per one-seat district) we should observe the same decrease in f on the entire country level, where the decrease is from $f(V)$ to $f(D)$, and on the level of a single district, where the decrease is from $f(V/D)$ to $f(1)$. Since $f(1)=0$, this yields

$$f(V/D) = f(V) - f(D),$$

an identity satisfied only if $f(V) = \log V$. This completes the derivation of

$$n(V,D) = f(V)/f(D) = \log V/\log D.$$

Notes

I would like to thank Philip Converse, Creel Froman, Bernard Grofman, Manfred Holler, Arend Lijphart, Howard Margolis, Frederick Reines, and Matthew Shugart for helpful comments.

1. For four parties of equal size (vote distribution of .25/.25/.25/.25), N is exactly 4, as one would want it to be. For a distribution of .45/.35/.19/.01 we obtain $N=2.77$, which tells us that there are only two major parties plus a third one of somewhat lesser importance. The Rae-Taylor fractionalization index F is connected to N through $N=1/(1-F)$.

2. For all curves shown in Figure 3 it was assumed that the aforementioned cube root relationship (Taagepera, 1972) between the number of assembly seats and the number of voters applies perfectly, so that equation (5) becomes $n=(3)^{1/M}$. Also, N was given the average value at the given district magnitude reported by Taagepera (1984) for worldwide data: $N=1.25(2+\log M)$. Despite these simplifications, the variety of curves obtained is considerable. Even more variety can be expected if one uses the actual values of N, D, and V for particular countries.

3. The following ranges were used: for single-seat districts, $M=1$; for low magnitude, multi-seat districts, $2<M<4$; for medium magnitude districts, $5<M<9$; for high magnitude districts, $10<M<20$; and for very high magnitude districts, $M>100$. There are no cases with M between 20 and 100. For the low M, Japan is the only case with simple allocation rules available. For the very high M, the number of parties no longer is expected to matter, and the case with the highest M available (the Netherlands) was chosen.

4. The only omission was post-1970 Austria, where the electoral system shifted from $M=6$ to $M=19$. The U.S. House data missing in Nohlen (1978) were added, using approximately the same time span.

5. Parties with less than 2.5% of votes were not plotted. For such very small parties, wide random fluctuations in advantage ratio can be expected for any electoral system, because seats come in integer numbers only. If a party with 1.5% of the votes obtains one seat in an assembly of 200 seats, its advantage ratio is a low 0.67. If it should reach two seats, its A immediately jumps to a very high 1.33.

6. In Austria's case the effective M could be somewhat higher than the value $M=6$ used here, because the seats not allocated by district-wide quota of $q=100\%/(M+1)$ were allocated (by d'Hondt procedure) in four wider regions (Nohlen, 1978, p. 270). This could involve up to 20% of the total 165 seats—i.e., up to 8 seats per region, which thus would represent another district with $M=8$.

7. Among the high magnitude systems, Sweden and Luxembourg have a slightly lower N than Italy, but Italy was chosen as the low-N case because it is at the upper edge of the range of magnitudes. Like Austria, Italy allocates most of the seats by quota in the districts (with the definition of quota varying over time), but the remainder of votes are pooled for nationwide distribution, with some restrictions on parties which do not get any seats in the districts. The remainder has involved from 20 to 80 seats (Nohlen, 1978, p. 275)—i.e., an average of 50 out of a total of about 630. This is roughy equivalent to having one district of $M=50$, in addition to the regular districts where M varies from 1 to 36. The average value $M=19$ used here may slightly underestimate the effective average magnitude. In Finland, too, individual district magnitudes vary widely, from 1 to 24. The wider scatter of the Finnish data might be caused by the use of local electoral alliances, which have an unpredictable effect on the fortunes of lesser parties. In Italy the scatter may be reduced by the nationwide distribution of remainder seats.

8. Rae's empirical equation based on worldwide PR data predicts, of course, the same proportionality curve for all countries. Among those predicted by our model, this curve is closest to Finland's. No previous model seems to exist that would predict the shifts in proportionality profile curves from country to country.

9. Predictions of $A=\%S/\%V$ have been tested here, rather than predictions of the percentage of seats, because the percentage of seats is so highly correlated with the percentage of votes ($r^2=.981$ for our sample) that our model's explanatory performance ($r^2=.985$) does not stand out, although it outperforms all other approaches. Using $A=\%S/\%V$ instead of $\%S$ eliminates the PR component, thus accentuating the deviation from PR and making for a more stringent test.

10. This derivation is condensed from Taagepera (1973).

References

Coleman, James S. 1964. *Introduction to Mathematical Sociology*. New York: Free Press of Glencoe.

Cox, Edward F. 1972. *State and National Voting: 1910–1970*. Hamden, CT: Archon Books.

Grofman, Bernard. 1975. A Review of Macro-Election Systems. In Rudolf Wildenmann, ed., *Sozialwissenschaftliches Jahrbuch für Politik*, vol. 4. Munich: Günter Olzog Verlag, pp. 303–52.

Grofman, Bernard. 1983. Measures of Bias and Proportionality in Seats-Votes Relationships. *Political Methodology*, 9:295–327.

Gudgin, Graham, and Peter Taylor. 1979. *Seats, Votes, and the Spatial Organisation of Elections*. London: Pion Limited.

Kendall, M. G., and A. Stuart. 1950. The Law of Cubic Proportion in Election Results. *British Journal of Sociology*, 1:183–97.

Laakso, Markku. 1979. Should a Two-and-a-Half Law Replace the Cube Law in British Elections? *British Journal of Political Science*, 9:355–84.

Laakso, Markku, and Rein Taagepera. 1979. Effective Number of Parties: A Measure with Application to West Europe. *Comparative Political Studies*, 12:3–27.

Lijphart, Arend. 1984a. Advances in the Comparative Study of Electoral Systems. *World Politics*, 36:424–36.

Lijphart, Arend. 1984b. *Democracies: Patterns of*

Majoritarian and Consensus Government in Twenty-One Countries. New Haven and London: Yale University Press.

March, James G. 1957. Party Legislative Representation as a Function of Election Results. *Public Opinion Quarterly,* 21:521–24.

Nohlen, Dieter. 1978. *Wahlsysteme der Welt* [Electoral Systems of the World]. München: Piper.

Qualter, Terence H. 1968. Seats and Votes: An Application of the Cube Law to the Canadian Electoral System. *Canadian Journal of Political Science,* 1:336–44.

Quandt, Richard E. 1974. A Stochastic Model of Elections in Two-Party Systems. *Journal of the American Statistical Association,* 69:315–24.

Rae, Douglas W. 1971. *The Political Consequences of Electoral Laws.* New Haven, CT: Yale University Press.

Sankoff, David, and Koula Mellos. 1972. The Swing Ratio and Game Theory. *American Political Science Review,* 66:551–54.

Sartori, Giovanni. 1968. Representational Systems. *International Encyclopedia of the Social Sciences,* vol. 13. New York: Macmillan & Free Press.

Schrodt, Philip A. 1981. A Statistical Study of the Cube Law in Five Electoral Systems. *Political Methodology,* 7:31–54.

Taagepera, Rein. 1972. The Size of National Assemblies. *Social Science Research,* 1:385–401.

Taagepera, Rein. 1973. Seats and Votes: A Generalization of the Cube Law of Elections. *Social Science Research,* 2:257–75.

Taagepera, Rein. 1984. Effect of District Magnitude and Properties of Two-Seat Districts. In Arend Lijphart and Bernard Grofman, eds., *Choosing an Electoral System: Issues and Alternatives.* New York: Praeger, 91–101.

Taagepera, Rein, and Markku Laakso. 1980. Proportionality Profiles of West European Electoral Systems. *European Journal of Political Research,* 8:423–46.

Theil, Henri. 1969. The Desired Political Entropy. *American Political Science Review,* 63:521–25.

Theil, Henri. 1970. The Cube Law Revisited. *Journal of American Statistical Association,* 65:1213–19.

Tufte, Edward R. 1973. The Relationship Between Seats and Votes in Two-Party Systems. *American Political Science Review,* 67:540–47.

Vick, Alan J. 1979. Proportionality Profiles of the United States Congress: 1910–1970. Unpublished manuscript.

Rein Taagepera is Professor, School of Social Sciences, University of California-Irvine, Irvine, CA 92717.