Title
Predicting the Number of Parties: A Quantitative Model of Duverger's Mechanical Effect

Journal
American Political Science Review, 87(2)

ISSN
0003-0554

Authors
Taagepera, Rein
Shugart, Matthew Soberg

Publication Date
1993-06-01

DOI
10.2307/2939053

Peer reviewed

PREDICTING THE NUMBER OF PARTIES: A QUANTITATIVE MODEL OF DUVERGER'S MECHANICAL EFFECT

REIN TAAGEPERA *University of California, Irvine and Tartu University*
MATTHEW SOBERG SHUGART *University of California, San Diego*

The mechanical effect of electoral systems, identified by Maurice Duverger, can be estimated by means of a quantitative model. The model predicts the range within which the effective number of parties in a district should fall for a given magnitude (number of seats) of the district. At the national level, a related model predicts the effective number of parties based on the effective magnitude and the number of seats in the national assembly. The institutional variables considered—magnitude and assembly size—define a great portion of the structural constraints within which a given country's politics must take place. The model developed provides a good fit to data in spite of its having been developed from outrageously simple starting assumptions.

If one had to give a single number to characterize the politics of any country that employs competitive elections, it would be the number of parties active in its national assembly. This number would not tell the whole story, by any means; but it tells us more than any other single number or term could. The number of parties directly or indirectly affects other important aspects of how a political system functions, including how long its cabinets last if the system is parliamentary (Lijphart 1984) and how elections translate into "citizen control" of policymakers (Powell 1989). The number of parties is a most important feature in a country's politics and therefore in comparative studies also. We shall discuss the ways of defining this number operationally further on.

What determines the number of parties? History, present issues, and institutions all intervene. But if one had to give a single major factor, it would have to be the *district magnitude* (M), that is, the number of seats allocated in an electoral district (Rae 1967). The well-known Duverger rule says that one-seat districts tend to lead to two-party systems, while multiseat districts tend to go with multiparty systems (Duverger 1951, 1954; see also Riker 1982). One can be more precise, since even within the multiseat category, a larger M tends to go with a larger number of parties. But before we can attempt to define a relationship between district magnitude and the number of parties, we need to establish a means to measure the number of parties.

How To Measure the Number of Parties

In this work, we shall use the *effective number* of parties, rather than the actual number. This measure is defined as follows:

$$N = (\Sigma p_i^2)^{-1}, \qquad (1)$$

where p_i is the share of votes or seats won by the *i*th party (Laakso and Taagepera 1979). The effective number tells us the number of hypothetical equal-sized parties that would have the same effect on the fractionalization of the party system as have the actual varying sized parties.[1] The advantage of using the effective, rather than the actual, number of parties (or actual number of parties above some arbitrary cutoff) is that it establishes a nonarbitrary way to distinguish "significant" parties from less significant ones. The construction of the index is such that each party weights itself by being squared. Tiny parties contribute little to the index, while large parties contribute relatively more.[2]

The Empirical Relationship and Methods of Estimation

The average empirical relationship is expressed by "the generalized Duverger's rule":

$$N_s = 1.15(2 + \log M), \qquad (2)$$

where logM is the decimal logarithm of the district magnitude (see our earlier work, Taagepera and Shugart 1989, 142–55). Such empirical relationships are useful but not sufficient. We shall seek a quantitative model to explain why the average number of parties at a given M tends to be precisely that large and not much smaller or larger. Such a model would capture the outcome of how a nation's politics are filtered through one very important political institution, namely, the number of seats for which politicians compete. Qualitatively, this process is the Duverger *mechanical effect* of district magnitude. With M = 1, only one party can win; and this has a strong tendency to impel the political forces to conglomerate into two large parties in order to have some chance of winning. As M increases, more parties can possibly win seats in the given district; hence, more parties can afford to run.

In building our model, we shall seek an explanation grounded in something more firm than simply an observation that such and such an equation is the best fit to the data. We seek to find a starting point—or set of starting points—that possess a certain deductive elegance. We do so by taking a variable whose upper and lower extreme possible values are easily established and by assuming that the average value between the extremes will prevail. Upon confronting such assumptions with data, the original estimates can be corrected if necessary. We prefer to make our corrections not through an "error" term or by expanding from bivariate to multivariate analysis but by going back to the original assumption and considering factors that might perturb it and lead to a different estimate. Thus, we use a mix of alternating deductive and inductive approaches, always trying to find the simplest possible model that appears to gain explanatory leverage. Our method is rather unconventional in the social sciences but is typical of disciplines as diverse as physics and sabermetrics (baseball analysis).[3] Our method appears to be especially appropriate for fields in which it is possible to define clear upper and lower limits to the value that certain variables can take and also, of course, where variables already come in ready quantitative form (e.g., parties and seats, mass and energy, runs and wins). The method, as applied to electoral analysis, thus produces results that many may consider mechanical, with no specification of actual social processes included. Indeed, far from being a weakness of our method, such a feature is its principal strength.

We are, after all, seeking to quantify the mechanical effect of the political institution called district magnitude; therefore, one should not be surprised to find that the enterprise itself turns out to be mechanical. It is mechanical in the sense of accounting not for political issues, personalities, or culture but only for the limits imposed on such features by institutional structures. Institutions shape what politicians prefer and, even more, how they pursue their preferences. Our dependent variable, the number of parties, shows the outcome of the aggregation of politicians' pursuits. Furthermore, our theory is limited to the number of parties in the national assembly, as distinct from parties participating in elections. The Duverger mechanical effect deals only with the parliamentary parties. The indirect impact on electoral parties—the Duverger "psychological effect," as discussed by Reed (1990)—is outside our scope here.

Proceeding from the simple and basic to the more complex, we shall first consider the seats parties win in a single electoral district. We can then proceed to the nationwide combination of the district-level effects, also keeping in mind the supradistrict effects of adjustment seats and thresholds. District-level effects of electoral rules have largely been neglected. (Major exceptions are Cox 1990; Katz 1980; Reed 1990; and Sartori 1986). Yet one cannot hope to explain the more complex nationwide phenomena without first elucidating the relatively simple scene in a single district.

THE EFFECTIVE NUMBER OF SEAT-WINNING PARTIES IN A DISTRICT

It is now time to begin building a model of the effective number of seat-winning parties in a district, which we shall designate N'_s. (Symbols without the prime will indicate the corresponding quantity at the nationwide level.) In building such a model, we rely on three basic equations, defining relations among several variables and then combining these equations into one that expresses the effective number of seat-winning parties in a district in terms of the magnitude of the district. The steps are

1. Estimate N'_s in terms of the seat share of the largest party.
2. Estimate the actual (not effective) number of parties winning seats in a district.
3. Estimate the relationship between the actual number of seat-winning parties and the seat share of the largest of these parties.

The logic behind each of these steps is simple. The end result is to get an estimate of N'_s in terms of M. We know that N'_s is (by definition) an estimate of how many "significant" parties there are in a given distribution of seats, so the most useful way to estimate N'_s in terms of a single observable (or estimatable) variable is to do so on the basis of the seat share of the most "significant" of all the parties—the largest one. This is step 1. Step 2 is independent of the first and allows us to get closer to an estimate of the seat share of the largest party, which we need if step 1 is to lead us to our goal. Step 2 is necessary because, while we could not easily arrive at a deductive estimate for the *effective* number of parties in a district, for the *actual* number of parties, such an estimate is simple (as we shall see). Having that estimate, then, lets us get an estimate for the seat share of the largest party (step 3). The steps can then be algebraically brought together to give us an estimate of N'_s in terms of M, for which we shall provide a graphical test. In order to keep the exposition accessible to a wider political science audience, not all the details of these building-block steps will be shown.

Step 1. Relation of the Effective Number of Seat-winning Parties to the Seat Share of the Largest Party

The value of N'_s is to an appreciable extent determined by the share of the largest component. When the other parties are extremely splintered, $N'_s = 1/s_1'^2$. When all the other parties are as large as the largest party, $N'_s = 1/s'_1$. Let us test for the possibility that the actual values will be around the geometrical mean of these extremes:[4]

$$N'_s = s_1'^{-1.5}. \quad (3)$$

A graph of this relationship (not shown) gave satisfactory agreement with data but revealed that a best-fitting line would have given an estimate of the values of N'_s about 30% lower. The test of whether this is adequate will come with the confrontation of the final model with data. But first, we go to step 2.

Step 2. Relation Between District Magnitude and the Actual Number of Seat-winning Parties in a District

In a district of M seats, the minimum number of parties is 1 (when one party wins all the seats) and the maximum is M (when each seat is won by a different party). We expect there to be an average number of actual seat-winning parties lying somewhere between 1 and M, so let us estimate the actual number of seat-winning parties in a district, p', using the geometrical mean:

$$p' = M^{.5} \qquad (4)$$

This estimate produced a reasonably good fit but could have been better. A correction can be made, based upon the notion that a given district usually does not stand alone; rather, it is one of several districts from which the members of the larger parliament are elected.

Correcting for National Politics. The model can be improved by observing that there is a difference between a district of M = 20 within a larger country and a small country with an assembly of 20 seats, all elected within a single nationwide district. In the latter case, all parties are generated within the district. In the former case, some nationwide parties with no ready constituency within that particular district may still bring in funds and talent from the outside and possibly win a seat. It is the difference between a closed system (where the model of equation 4 may apply) and a more complex, open system. Our correction factor brings in the number of seats (S) in the national assembly.

The effect of assembly size S must enter in the form of the ratio S/M, since it must vanish when the entire country is made one single district (M = S). To satisfy the conditions that $M > p' > M^{.5}$ at any district magnitude, and $p' = M^{.5}$ at M = S, the expression for p' must have the form

$$p' = M^{k}, \qquad (5)$$

where $k = 1/[1 + (M/S)^n]$, n being a constant, as yet unspecified. Furthermore, p' at any M < S must not be higher than it is at M = S. This condition would be violated, if we set n = 1. The largest acceptable value of n would be one that makes p' almost reach the value of $p'(S)$ at large M < S; that means a value of n that makes the slope of the curve zero ($dp'/dM = 0$) for M = S. Taking the derivative of p' in the equation above and setting it to zero at M = S yields $n = 2/\ln S$. This is the maximum value n could take, corresponding to the maximum possible effect of nationwide politics on the given district. The minimum possible effect is, of course, no effect, which corresponds to n = 0 (and thus $p' = M^{.5}$ at any M). In the absence of any further information, the expectation value of n is taken as the mean of the extremes, that is, $n = 1/\ln S$. Now equation 6 becomes

$$k = 1/[1 + (M/S)^{1/\ln S}]. \qquad (6)$$

For a single nationwide district (M = S), this reduces itself to k = .5 and thus equation 2. For M < 30, the value of p' is very little affected by the precise value of S, as long as it is larger than 60. The median assembly size for independent countries is around 150, and the usual range is 50–500. The value S = 148 corresponds to the round figure of $1/\ln S = .200$. Because $1/150^{.200} = .37$, the result is

$$k = 1/[1 + .37M^{.2}]. \qquad (7)$$

This correction for the effects of nationwide politics significantly improves the fit to our data. In our graphical presentation of the relationship between M and N'_s further on, we present equations with both the exponents .5 and k as defined in equation 7.

Step 3. Relation Between the Actual Number of Seat-winning Parties and the Seat Share of the Largest Party

The final independent link in the chain leading to an expression that links N'_s and M is to estimate the relationship between the number of actual seat-winning parties and the seat share of the largest party, s'_1. The average fractional seat share of the seat-winning parties must be $1/p'$. The fractional share going to the party with the most seats in the given district, s'_1, must be at least this average. The upper limit on s'_1 is 1, when that party wins all the seats. In the absence of other knowledge, the expected average value is again the geometric mean:

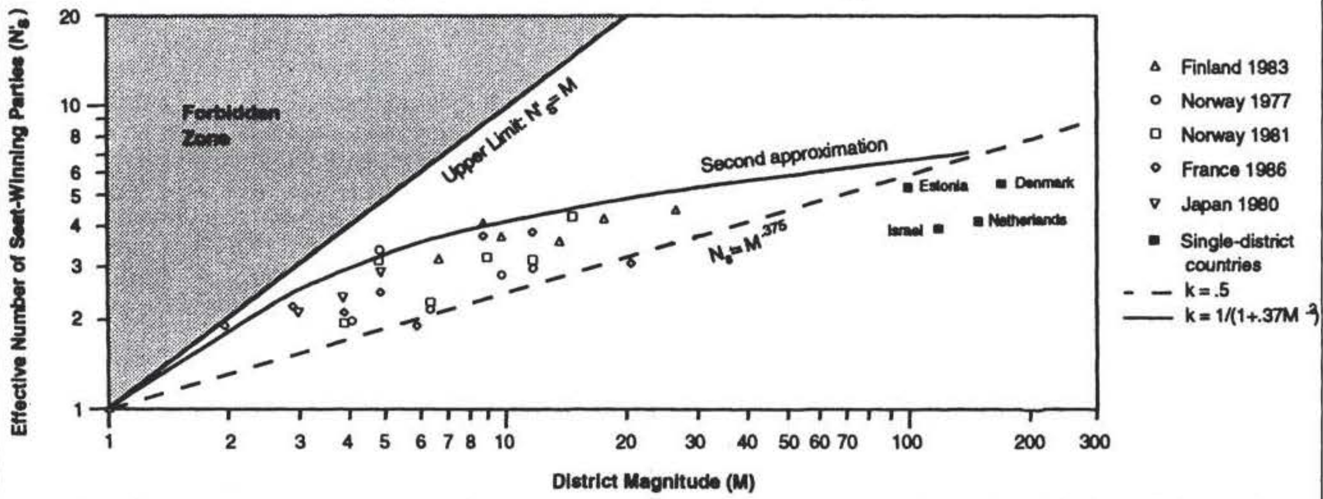
$$s'_1 = 1/p'^{.5}. \qquad (8)$$

This estimate is completely independent of the relationship between p' and M. Indeed, even if p' and M were not correlated, equation 8 would still be independently derivable. It should be noted that the largest party in the given district need not be the largest nationwide.

A graphical test of equation 8 showed that the estimate captured the general trend but that data points tended to fall below the curve. Even so, any agreement is quite gratifying, given the simplicity of the assumption that led to the estimation. As before, this simple first-approximation model (equation 8) can be improved upon. The upper limit on the seat share of the largest party is actually *less* than one (rather than one, as is assumed in the first approximation).[5] The test of whether the agreement of either the original model or the second approximation is satisfactory or not will come when we bring together all the building blocks and test our resulting estimation of N'_s based on M against actual data.

FIGURE 1

Effective Number of Seat-winning Parties in District Versus District Magnitude



Combining Steps: The Relation between District Magnitude and the Effective Number of Seat-winning Parties in a District

It is now time to combine all the previous steps and express N'_s as a function of district magnitude. The first-approximation model[6] leads to

$$N'_s = M^{.375}. \quad (9)$$

Figure 1 shows the curve that results from this equation along with data. The actual data points tend to be above this line.

Let us now introduce our corrections. The curve labeled "Second approximation" takes account of the corrections discussed above for both p and s_1.[7] Most actual data points fall between the two curves, suggesting overcorrection or an effect of *two-party competition* that will be discussed later.

Discussion of Results: The Overall Pattern. Two objections may be raised at this point: (1) such a probability-based model completely ignores the political processes that lead, through elections, to the winning of seats; and (2) it does not even fit the data very well. The two objections actually neutralize each other. If all the data points fell perfectly on the model-predicted line, then, indeed, the political process would be out of the picture. Only institutions would matter—indeed, only one institution, the district magnitude. The very scatter of the points in the actual picture indicates the workings of the various political processes. The range of possible outcomes is nonetheless restricted. No data could be in the "forbidden zone" in which $N'_s > M$, but all the rest of the space of the figure is fair game. Yet all the data are found in a narrow zone. This should not surprise us. Political processes do not take place in a vacuum. They and their outcomes face various restrictions, including logical constraints, of which statistical probabilities are an aspect. The political institutions (particularly the district magnitude) are a critical element providing these constraints on possible political outcomes. Our task here is to define the average pattern and find out the extent of the actual playroom left for political processes.

The achievements of the simple models shown in Figure 1 should not be underestimated. All the data points could conceivably have fallen much above or much below the predicted line. That this is not the case represents success, for a first approximation.

Poor Fit of the Single-District Countries. The lower values of N'_s for single-district countries (those in which all seats are allocated in one nationwide district) might imply that for such systems, there is some ceiling for a given M above which the effective number of parties does not rise. A reason for such a reduction below the theoretical prediction could be that most of a country's ambitious politicians (those who desire to sit in cabinets or at least have policy influence) prefer to join one of the larger parties, rather than a very small party. They do so with a greater likelihood than would be predicted on the basis of opportunities for small parties provided by high M. The larger parties' seat shares are thus augmented, so that N'_s is lessened. Thus, although tiny parties based on narrow interests can and do form and win seats, the effective number of parties (high though it may be relative to other countries) is actually *lower* in countries such as Israel and the Netherlands than what M would theoretically allow for. When we get to an estimation of the nationwide N_s including countries with multiple districts, the fit for the single-district countries improves considerably. We now turn to the nationwide scene.

THE NATIONWIDE NUMBER OF SEAT-WINNING PARTIES

Our process of building a model for the nationwide effective number of seat-winning parties (N_s) closely parallels the process used at the district level. Indeed, for one of the building blocks, it matters not at all whether the quantity being measured is district- or national-level. For the relation between the seat share of the largest party (s_1) and the number of actual seat-winning parties, equation 8 holds. The other two steps, however, require some fine-tuning at the national level.

Relation of the Nationwide Effective Number of Seat-winning Parties to the Seat Share of the Largest Party

The basic reasoning for the relationship at the national level between the effective number of parties and the seat share of the largest parties is exactly the same as for a single district for which we derived equation 3: $N_s = s_1^{-1.5}$. However, at the nationwide level another consideration enters, namely, special competition between the two largest parties. The nationwide vote shares of the two largest parties have lately been extremely close in a number of countries (West Germany, New Zealand, Israel, and Switzerland, in their latest election according to Mackie and Rose 1991); and seat shares have also been fairly close. The existence of one large party strongly encourages the opposition to unite or else remain almost permanently in limbo. This can be seen even in Israel despite electoral rules that almost invite the creation of further small parties. Because of such pressures, the second party's vote share has lately been more than one-half of the largest party's in the wide majority of stable democracies, the only exceptions being Denmark, India, Japan, and Sweden. When we consider the effects of two-party competition,[8] we get the following equation for $s_1 < .5$:

$$N_s = .85 s_1^{-1.5} \quad [s_1 < .5]. \tag{10}$$

This correction is not expected to apply to individual districts. The two major parties may compete neck-and-neck for the nationwide total of seats, but why would they throw scarce resources into hopeless individual districts just to try to achieve district-level parity of seats? However, Figure 1 suggests that parties may do exactly that; most data points in Figure 1 would fall between the analogues of the equations developed here to correct for two-party competition at the national level.

Relation Between Magnitude and the Nationwide Actual Number of Parties

We cannot consider M at the national level in the same way as we do at the district level simply because M, *district* magnitude, is by definition a district-level phenomenon. Except for those systems in which all seats are allocated in one district, the national level forces us to come to grips with the aggregation of several districts. Our reasoning in developing nationwide expressions assumes districts with equal magnitude with all seats allocated in these districts (i.e., no seats allocated in a regional or nationwide tier). Almost no country in the world (with $M > 1$) fits this description; but let us solve the simple case first, then see whether the resulting expressions will work for the more complex systems.

The actual nationwide p is likely to be higher than district-level $p' = M^k$ (equation 5), as some parties win seats in some districts but not in others. If the entire country were made a single electoral district of magnitude S, the number of seat-winning parties would be S^k. Thus, $M^k < p < S^k$; and we can again expect the actual values to be around the geometrical average of the possible extreme values:[9] $p = (SM)^{.25}$. At the first glance, this equation suggests that district magnitude and the total size of the assembly are equally important in determining the nationwide number of seat-winning parties; but this is not quite so. The range for S is from about 13 to 650, a ratio of 1:50, while M ranges from 1 to 150, a ratio of 1:150. Hence, M influences the number of parties appreciably more—so much more, in fact, that the effect of S has apparently not been noticed previously.

Combining Steps: Estimating the Nationwide Effective Number of Parties Based on Magnitude

We can now combine our three steps at the national level just as we did at the district level and arrive at an expression for the nationwide effective number of parties. The resulting expression is[10]

$$N_s = .85(SM)^{3/16} \tag{11}$$

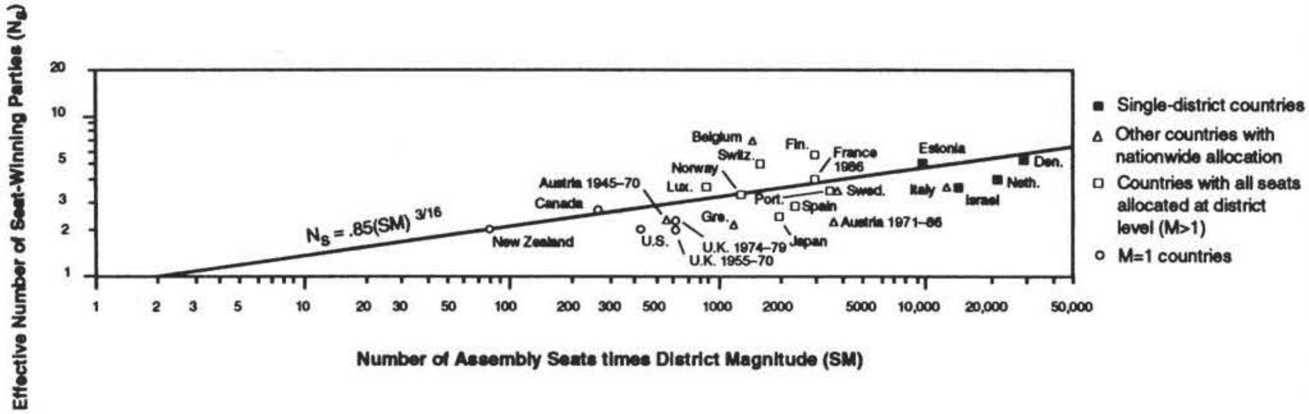
Figure 2 shows basic agreement of this estimation with data. This applies to countries with large assemblies and $M = 1$, as well as to those with smaller assemblies but multiseat districts.

Four different types of countries are distinguished in this graph: (1) nationwide $M > 1$ (i.e., a single district), with a threshold of no more than 2%; (2) $M > 1$, with appreciable supradistrict seat allocation and/or high thresholds; (3) $M > 1$, with seat allocation purely within districts; and (4) $M = 1$ plurality.[11] For the second category, the "effective magnitude" differs considerably from the average district magnitude, which we use for the other cases in which district magnitude varies within the system (see Taagepera and Shugart 1989, 126–41). For the cases with supradistrict allocation and/or thresholds, the agreement of data with estimates based on our simplifying assumptions can be expected to be reduced. However, this does not seem to be the case.

It should be kept in mind that within the same electoral system, the individual outcomes can vary widely. In Finland, for example, the electoral rules have remained the same ever since 1907 (including S and M); and on the average, the theory predicts

FIGURE 2

Nationwide Effective Number of Seat-winning Parties Versus Product of the Number of Assembly Seats and District Magnitude



$s_1 = .37$ and $N_s = 3.7$. The actual s_1 in Finland has ranged from .25 in 1958 to .51 in 1917, while N_s has ranged from 3.0 in 1917 to 5.5 in 1970. Over the time period taken into account here (1966–87), N_s in Finland has been very close to that predicted by equation 11. Approaches to political science such as that used here imply that over time, party leaders and voters should learn how to play the game as it is established by a country's institutions (Reed 1990). The result of such learning is an adjustment of outcomes, such as N_s, to institutions, such as SM. Figure 2 tells us that two institutions—district magnitude and assembly size—rather remarkably account for the effective number of parties in a country. Only a few countries appear as truly deviant. One is Austria after 1971 (although not before then, as we shall discuss); and another is Belgium. Both of those cases deviate by more than 40%. In the case of Belgium, the number of parties has not always been so great; and it could readjust downward once again. Alternatively, parties could choose to reform the electoral system by increasing magnitude, the assembly size, or both, as often happens when the number of parties rises for some time (Shugart 1992).

We have yet to express the effective number of parties nationwide as a function of magnitude alone so as to compare it to the empirical fit in equation 2. Apart from the constraint $S > M$, there is no logical or empirical connection between M and S. As observed earlier, the median S is around 150. Using this approximation to eliminate S from equation 11 leads to equation 12, which is graphed in Figure 3:

$$N_s = 2.15\ M^{3/16} \quad [M < 150]. \qquad \textbf{(12)}$$

For $M = 1$ countries, this equation predicts $N_s = 2.15$, while equation 2 yields 2.3. Throughout the range $1 < M < 40$, the theoretical equation yields values within $\pm.15$ of the empirical equation 2. For larger magnitudes, the theoretical equation yields larger values: for $M = 100$, $N_s = 5.1$, as compared to 4.6 given by equation 2. Both equations express the general trend to an equal degree, and equation 12 can be said to give theoretical explanation and validation to the empirical equation 2, while modifying it to some degree.

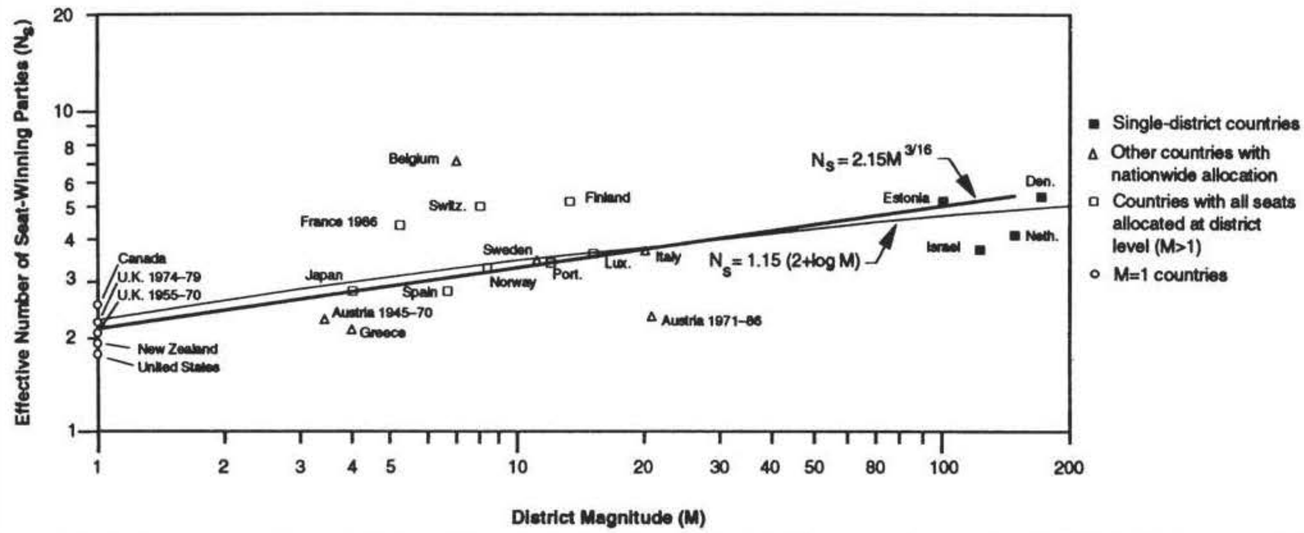
However, there is no need to neglect the role of S. In most cases, a country's assembly size is set first—usually within a narrow range approximated by the cube root of population (Taagepera and Shugart 1989). Once the size of the assembly is set, the magnitude usually follows as a result of how many administrative subdivisions there are—with the exceptions of countries using $M = 1$ or an electoral formula that almost requires small magnitudes.[12] When politicians desire to represent more parties proportionally, they must either increase S (automatically increasing M) or else decide to allocate some seats in a nationwide compensatory district. Thus, the choices of S and M are not entirely independent of one another; but neither does the choice of one value for one variable absolutely restrict the choice on the other variable (unless constrained by, e.g., requiring that all seats be allocated within districts *and* that each preexisting territorial subdivision must form its own district). Compared to equations 2 or 12, equation 11, accounting for S, improves the correlation and is preferable from a theory viewpoint.[13]

DUVERGER'S MECHANICAL EFFECT: THE QUANTITATIVE MODEL

The Duverger mechanical effect connects the number of seat-winning parties to the district magnitude. For a long time, the theory of this effect has remained on a qualitative level. We have now presented and tested the individual building blocks—or, rather, links in a quantitative concatenation ranging all the way from district magnitude M to the effective number of seat-winning parties on the district and the

FIGURE 3

Nationwide Effective Number of Seat-winning Parties Versus District Magnitude



nationwide levels. Some of these links are firmer than others. We can join them, so as to express everything in terms of M only. While doing so, every link introduces its share of error. In the case of N'_s (district level), the concatenation of simple elements involves three links. For N_s (national level), more links are involved, including the very coarse approximation $S = 150$ if one wants to eliminate S.

We would expect the agreement to become increasingly worse, as we add steps. If any similarity between the final model and the data is found, then the model must be considered quite robust, indeed. This is true, in particular, for the nationwide number of parties, where the average district magnitude is used, although a few very large districts may affect the outcome and supradistrict seat allocation occurs in many countries. In this perspective, ending up with results that agree at all with the general pattern of N_s versus SM must be considered a success.

For the district-level number of parties, Figure 1 suggests that the district magnitude determines the range of outcomes within a factor of two for most individual elections and within a much narrower zone for averages of many elections in the same country. People are certainly free to vote into office many more or many fewer parties in a given district, but it very rarely happens. The implication is, then, that the institutions reviewed here (mainly M, but also S) provide predictable constraints on outcomes. Politicians and voters adjust their behavior in a way that ensures that outcomes accord well with our quantitative predictions.

Where do we find significant deviations from predicted values? We have discussed the cases in which all seats are allocated in one district. These cases deviated on N'_s, although not on p′. We suggested also that national politics impinges on individual districts in those countries with multiple districts, leading us to the correction represented by equation 5, taking account of the size of the national assembly. What we find is that national politics impinges upon an individual district even when it is the only district. In work not shown here, we found that for these countries, the actual number of parties (p′) is about what equation 4 predicts; but the effective number (N'_s) is less than what equation 9 predicts. This suggests a novel conclusion, namely, that the incentives to control executive office in these parliamentary systems with nationwide allocation leads to *less* extreme fragmentation than the institutional design alone would imply. That a national-level factor is important for these countries, as well as for the rest of the sample, is suggested by the greater fit for these cases to the models derived for nationwide results (equations 11 and 12), which introduce assembly size.[14]

Two cases generally regarded as deviant in the literature on Duverger's law (as reviewed by Riker 1982) deserve attention here. They are Austria and Canada, which will be discussed alongside other contrasting cases. Austria is often thought to be exceptional, because it uses proportional representation yet has only two important parties. The sociological dominance of the two big parties' organizations has been invoked as an explanation (Riker 1982). If we look to the period since 1971, indeed, Austria looks deviant. However, before 1971, Austria's electoral system was based on small districts with restricted access to allocation of nationwide compensatory seats (Kitzinger 1959), such that the effective magnitude was about 3.5 (Taagepera and Shugart 1989, 138). With such a low magnitude, a low number of parties can be expected. The effective magnitude was increased to about 20 in 1971. If

Austria's N_s remains low for a long period of time, then we would indeed need an extrainstitutional explanation. However, recent elections have seen a rise in N_s, which may presage a gradual adjustment of the party system to the new electoral rules.

In Canada, the number of parties is often claimed to be "too high" for Duverger's law to be valid. Riker (1982) suggests federalism as the explanation: parties that are based in provinces also run nationally; thus, N_s is appreciably greater than 2.0 despite $M = 1$. However, Canada's N_s is almost precisely what it is predicted to be when S is taken into account. On the other hand, the United Kingdom, almost always the standard example confirming Duverger's law, is actually below predictions (although not enough to categorize it as deviant). Thus, rather than invoking federalism to explain a relatively high number of parties in Canada, perhaps we should invoke *lack of federalism* to explain the relatively low number of parties in the United Kingdom,[15] much as presidentialism can explain why the United States is also low.[16] Federalism need not be invoked for Canada unless N_s rises and stabilizes without a prior increase in S.

Besides the case of Austria since 1971 (already discussed), Belgium is the only case that deviates substantially from predictions even when the assembly size is taken into account. Figure 3 shows Finland and Switzerland also far off that theoretical curve, which does not account for a country's particular assembly size.[17] These cases have "too many" parties for their electoral rules. Yet no one ever calls these countries deviant: they use proportional representation, so they are *supposed* to have a lot of parties! Since none of these countries has always had such high values of N_s, it is possible that either their values of N_s will at some time decline once again or else the institutions will be reformed. If either event resulted, it would indicate a long-term adjustment of behavior and institutions to one another. The models here do not, however, allow us to predict when such adjustment would occur or, indeed, even if it will occur. This discussion has suggested that a promising next step would be longitudinal analysis of individual countries, to look more directly at the process of adaptation (see Reed 1990; Shugart 1992).

CONCLUSIONS

What have we achieved? We have asked simple questions like, How many parties might one expect to see in an electoral district? These simple questions have apparently not been asked previously, largely because attention has been fixed on the national level. The national level is more important, but it is not conducive to model building (without looking first into district-level phenomena); hence, analysts were left with primarily empirical expressions. On the district level, simple models can be constructed and then expanded to the nationwide. The simplest is to establish the upper and lower boundaries of a variable and assume that average values in between these boundaries prevail. This approach has been quite successful: the model agrees with the data surprisingly well.

The interaction between empirical work and model building is a two-way street. In the present case, empirical results at time urged a search for a theoretical explanation, while at other times theoretical advances suggested that unsuspected relationships be tested empirically.

There is a broader question about this exercise that is worth pondering. That is the question of what these models really tell us about politics, as opposed to simple numerical descriptions of average political outcomes. These models should not be construed as apolitical, for two reasons. First, at least as far back as Duverger (1951, 1954), the connection between electoral systems (mainly defined by their district magnitude) and the number of parties represented in assemblies was recognized as "mechanical." Since political institutions, including district magnitude, are a machine that converts the actions of politicians and voters into political outcomes, it is hardly surprising that models of the effect should themselves appear "mechanical" in form.

Second, we have not offered any models of how individual politicians calculate costs and benefits, as do some rational-choice theorists in political science. Those are micro concerns and involve forays into what Duverger termed the "psychological" effect. Our models suggest that the aggregation of individual political actors' calculations, made in part according to the institutional context in which winning or losing will be determined, occurs in predictable fashion. Perhaps some day we may have models of how this aggregation itself takes place. The remarkable regularity across countries in their political outcomes (at least in terms of the number of parties) according to the form of their political institutions gives us some hope that we have taken a step in the direction of such more generalizable models of politics.

Nor have we advanced herein a theory of how electoral rules are chosen, although we have suggested that where the existing number of parties is "too high" for the existing magnitude, a "psychologically" induced reduction in the number of parties is not the only possible outcome. Indeed, electoral reform might result, increasing magnitude. Further theoretical and empirical work should address itself to the interaction among the mechanical effect, psychological effect, and choice of electoral rules. We have made a preliminary report on the first part of this link.

Notes

1. Indeed, N is merely a transformation of Rae's index of fractionalization: $N = 1/(1 - F)$. We prefer N because it is easier to visualize in concrete terms. For example, $N = 2.84$ suggests that there are more than two, but definitely less than

three, major parties, while $F = .648$ less obviously corresponds to the number of components in the system.

2. N is not the only measure that could be used. See Laakso and Taagepera 1979 and Taagepera and Shugart 1989 for a discussion.

3. Sabermetrics is the quantitative analysis of regularities in aggregate performance in baseball. Although it has (yet) to claim a single university department of its own, it is indeed a "scientific" discipline with sophisticated methods seldom seen by (or of any interest to) most fans or sportwriters. See James 1984, 9–24, and *Baseball Abstract* in general.

4. The geometrical mean is computed by multiplying the values of the upper and lower extremes, then taking the square root of the product. It may be thought of as an arithmetic mean in log–log form, which we have used here because, in order for an assumption of linearity to work on these data, one must first transform the data through decimal logarithms.

5. The reason is that if more than one party wins seats in a district, the largest party obviously can not win 100% of the seats. In a district with M seats, if p′ parties win at least one seat each, then the largest party cannot win M seats but only $M - p' + 1$ at the most. Its maximum share of the seats then is $(M - p' + 1)/M$, rather than 1. Since M can, in principle, be expressed in terms of p′—by reversing the directions of equations 5 and 8—s_1' remains a function of p′ alone, as it was in equation 8. However, the algebraic form of the relationship is quite complex. Rather than finding it, it is more fruitful to determine the second-approximation curve by calculating various points on it. For instance, for $M = 9$, equations 5 and 8 yield $p' = 4.04$ and maximum s_1' is $(9 - 4.04 + 1)/9 = .66$ (rather than 1). The geometric mean of the extremes becomes $s_1' = (.66/4.04)^{.5} = .40$ (rather than .50, as given by equation 8). Calculations of this sort yield a "second approximation" that visibly improved the agreement with data.

6. This first approximation is derived from the aforementioned building blocks, as follows. Our goal is to estimate N_s as a function of M. The one variable that we were able to estimate deductively as a function of M was p, so we start there: $p = M^{.5}$. We also estimated that $s_1 = 1/(p^{.5})$; therefore, $s_1 = 1/[(M^{.5})^{.5}] = 1/(M^{.25}) = M^{-.25}$. Now we return to $N_s = s_1^{-1.5}$ and plug in our estimate of s_1 in terms of M: $N_s = (M^{-.25})^{-1.5} = M^{.375}$.

7. Equation 9 may be respecified as $N_s = M^{.75k}$, where k is defined as .5 in the first approximation (derived from equation 4) or according to equation 7 in the second approximation.

8. The derivation of this estimation is as follows. In the case of perfect two-party competition ($s_2 = s_1$), the upper limit on N_s is no longer $1/s_1^2$ but $1/2s_1^2$, while the lower limit remains $1/s_1$. The geometric mean of the extremes becomes

$$N_s = 2^{-.5}\, s_1^{-1.5} = .71 s_1^{-1.5}\ [s_1 < .5]. \qquad \textbf{(13)}$$

A graphical analysis showed that, for $s_1 < .5$, the actual data points fall largely in the zone delineated by equation 13 and equation 3, that is, in the zone between full two-party competition and no two-party competition. The average pattern for $s_1 < .5$ is the mean of the two pure cases: $N_s = .85 s_1^{-1.5}$. For $s_1 > .5$, the second-largest party must be smaller than the largest one, and equation 13 is no longer valid.

9. We use $k = .5$ instead of the more complex form given by equation 7. The latter expresses the effect of nationwide politics on the district, and using it here would mean correcting the nationwide politics for itself.

10. The algebraic derivation proceeds as follows. We have estimated that $s_1 = (1/p^{.5})$ and that $p = (SM)^{.25}$. Therefore, $s_1 = 1/[(SM)^{.25}]^{.5} = (SM)^{-.125}$. Plugging this into equation 10, we get $N_s = .85[(SM)^{-.125}]^{-1.5} = .85(SM)^{.1875} = .85(SM)^{3/16}$.

11. The countries and time periods used, by category identified in the text, are (1) Denmark, Estonia (1923–32), Israel, and the Netherlands; (2) Austria, Belgium, Greece, Italy, and Sweden; (3) Finland, France (1986), Japan, Luxembourg, Norway, Portugal, Spain, and Switzerland; (4) Canada, New Zealand, the United Kingdom (1955–70, 1974–79), and the United States. Median values are used from the last seven elections listed in Mackie and Rose 1991 except as indicated or where a change of electoral system had occurred during the period.

12. The single transferable vote (as in Ireland) and single nontransferable vote (as in Japan) are rarely used with districts with magnitudes much greater than six.

13. The previous empirical expression (equation 2) did not include S because its effect could not be observed easily, overshadowed as it was by the effect of M.

14. Since all the nationwide cases use S of around 150, the fit to the model that introduces this average value (equation 11) is no worse than that of the model that uses each country's specific value of S (equation 12).

15. If the United Kingdom ever implements a home rule parliament for Scotland, we will have a test. With a local parliament as a "prize," perhaps the Scottish National party would perform better. If this were carried over into national elections, N_s would increase somewhat.

16. Presidential systems can have a considerable and varied effect on the number of parties in the assembly (Epstein 1967; Shugart 1988); and the present simple theory does not take this effect into account. See Shugart and Carey 1992, 230–37, 293–300 for related models that do account for the "presidential difference."

17. France in 1986 looks deviant in Figure 3; but when its large S is accounted for (Figure 2), the model accounts almost perfectly for the effective number of parties in that system.

References

Cox, Gary. 1990. "Centripetal and Centrifugal Incentives in Electoral Systems." *American Journal of Political Science* 34: 903–35.

Duverger, Maurice. 1951. *Les partis politiques*. Paris: Le Seuil.

Duverger, Maurice. 1954. *Political Parties: Their Organization and Activity in the Modern State*. Trans. Barbara and Robert North. New York: Wiley.

Epstein, Leon. 1967. *Political Parties in Western Democracies*. New Brunswick: Transaction Books.

James, Bill. 1984. *Baseball Abstract*. New York: Ballantine.

Katz, Richard S. 1980. *A Theory of Parties and Electoral Systems*. Baltimore: Johns Hopkins University Press.

Kitzinger, Uwe. 1959. "The Austrian Electoral System." *Parliamentary Affairs* 12:392–404.

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

Lijphart, Arend. 1990. "The Political Consequences of Electoral Laws, 1945–85." *American Political Science Review* 84: 481–96.

Mackie, Thomas, and Richard Rose. 1991. *The International Almanac of Electoral History*. Washington: Congressional Quarterly.

Powell, G. Bingham, Jr. 1989. "Constitutional Design and Citizen Electoral Control." *Journal of Theoretical Politics* 1:107–30.

Rae, Douglas W. 1967. *The Political Consequences of Electoral Laws*. New Haven: Yale University Press.

Reed, Steven R. 1990. "Structure and Behaviour: Extending Duverger's Law to the Japanese Case." *British Journal of Political Science* 20:335–56.

Riker, William H. 1982. "The Two-Party System and Duverger's Law: An Essay on the History of Political Science." *American Political Science Review* 76:753–66.

Sartori, Giovanni. 1986. "The Influence of Electoral Systems: Faulty Laws or Faulty Method?" In *Electoral Laws and Their Political Consequences*, ed. Bernard Grofman and Arend Lijphart. New York: Agathon.

Shugart, Matthew S. 1988. *Duverger's Rule, District Magnitude, and Presidentialism*. Ph.D. diss., University of California, Irvine.
Shugart, Matthew S. 1992. "Electoral Reform in Systems of Proportional Representation." *European Journal of Political Research* 21:207–24.
Shugart, Matthew S., and John M. Carey. 1992. *Presidents and Assemblies: Constitutional Design and Electoral Dynamics*. Cambridge: Cambridge University Press.
Taagepera, Rein, and Matthew S. Shugart. 1989. *Seats and Votes: The Effects and Determinants of Electoral Systems*. New Haven: Yale University Press.

Rein Taagepera is Professor of Political Science, University of California at Irvine, Irvine, CA 92717 and Dean of the School of Social Sciences, Tartu University, Estonia.

Matthew Soberg Shugart is Assistant Professor of Political Science, University of California at San Diego, La Jolla, CA 92093-0519.