Title
Limiting Frames of Political Games: Logical Quantitative Models of Size, Growth and Distribution

Author
Taagepera, Rein

Publication Date
2002-05-01

CSD Center for the Study of Democracy
An Organized Research Unit
University of California, Irvine
www.democ.uci.edu

Politics takes place in time and space -- both the immutable physical space and the institutional space that politics can alter, but with much inertia. The size of polities matters for their functioning, as stressed early on by Robert Dahl and Edward Tufte in *Size and Democracy* (1973). Institutional size places constraints. For instance, in a 5-seat district at least one party and at most five can win seats. Within these bounds politics is not predetermined, but at and close to these bounds the limiting frame does restrict the political game.

Things take time. Sizes of populations, polities and markets change over time, but not instantaneously. Change rates over time impose further constraints on politics and on factors that affect politics. This is another limit.

The effects of such limiting frames have concerned me over 35 years, and it is time to give an overview of my contribution to this aspect of political science.[1] The purpose is twofold: Make my approaches and results more understandable (even to myself), and enable others to extend the use of this particular set of methods to further problems in political science and related fields.

The general focus of this endeavor has been modeling the impact of physical and institutional constraints on the political process and the society more generally. By modeling I mean *logical quantitative* models. These models go beyond empirical correlation analysis that is quantitative but does not specify logical connections. They extend logical reasoning to a degree of quantitativeness that goes beyond "When *x* increases, *y* must decrease". A logical *and* quantitative model is more specific: "When *x* changes by a certain amount, *y* must change by a specific corresponding amount, because of the following mechanisms or constraints". The resulting theoretical curves, y vs. x (or in many variables), can then be compared to empirical findings.

More specifically, the issues I have tried to model fall into the following four categories.

1) The size of countries, assemblies and electoral districts matters for their functioning. Exactly how does country size affect the size of its legislative assembly, its foreign trade/GNP ratio, and the sizes of cities?
2) Sizes of populations, countries, and defense budgets change over time -- growing, declining, interacting. How do more universal patterns of growth and duration enter those of empires and cabinet coalitions? How do social phenomena such as population explosion and hyperinflation proceed?
3) The number and size distribution of political parties is affected by institutional frameworks. According to which logical models?
4) Finally, conversion from people to representatives takes place in several forms. Popular votes translate into assembly seats for different parties. Populations of countries determine their seat shares in supranational assemblies. The conversion is usually less than perfectly proportional, under-representing the smaller parties, yet over-representing the smaller nations. What are the hidden rules of conversion?

Such are the questions that have puzzled me. My physics background has influenced the choice of topics tackled and the methods used. I have preferred issues where hard numerical data are available at least in principle -- square kilometers, populations, numbers of votes and seats, durations in months -- rather than percentages in opinion polls, game theoretical combinations and the like. I appreciate the value of the latter but personally prefer to focus on issues where logical quantitative models can be built and tested with hard data. It puts me outside the study of the core of the political process, but when the core processes are affected by outside constraints, then being outside the core is not the same as being marginal. Indeed, some cores may be disentangled by first addressing the external layers in a systematic way.

Some of my topics extend into related fields. Trade/GNP ratio belongs to political economy, the populations of cities within countries to geography, and world population growth to demography. I am not concerned with disciplinary borders. These issues bring understanding about the general methodology used and thus contribute to potential further applications in politics.

This overview is limited to those aspects of my work that are universal -- models that apply over time, space and cultural differences, to the extent the underlying assumptions hold. This is science in the strict timeless meaning of the term. Political science also includes painstaking collecting and processing of empirical data, as well as softer aspects: qualitative area studies, fashionable trends, and glorified journalism. I have done my share of it, including about 50 articles (some of them quite quantitative) and several books on small East European nations.[2] This is outside the scope of the review. So are numerous analyses of electoral systems that involve no model building as such.

I have received support and encouragement from many colleagues. The short list of thanks would have to include Harry Eckstein, Bernard Grofman, and Arend Lijphart. Going further back, my training in logical thinking began with my mother. When I guessed at 11 times 3 being 28, she did not say it was the wrong answer but built on my knowing that 10 times 3 just meant adding a zero. She said: "But how much is 10 times 3? Then how can 11 times 3 be less?" This is the moment when the 5-year-old realized that, beyond authoritative transmitted knowledge, there was such a thing as logical reasoning, in everyone's reach. And then there was this young physics instructor, Monsieur Flèche, at Lycée Mangin in Marrakech who drove me nuts by always muttering about things being actually more complex than he was supposed to tell us. I wanted simple answers. We both were right.

1. Lessons from Physics

A logical order would be to present first my general philosophy and method in some detail, pointing out its strong, weak and complementary aspects, compared to those now predominant in political science. Only thereafter would the approach be applied. Yet I start here with only short and incomplete methodological notes, because it's the results that matter. If my results awake interest in readers, then they might wish to look further into the method. If not, why bother?

The payoff of any method should be in expanding our explanatory power. Yet the very term "explanatory" needs explaining, because it has undergone a risky shift since the times that canned statistical analysis programs became widespread. Suppose a multifactor correlation analysis tying y to x_1 and x_2 leads to R^2=0.60. This would mean that x_1 and x_2 *account* for 60% of the variation in y. In the statistician's shorthand, they "explain" 60% of the variation, and both may be statistically significant. But throw in another variable x_3, and R^2 may go up while the previously significant x_1 may become statistically non-significant. Then what does x_1 "explain" in a substantive way?

The statistician cannot be faulted for using "explain" as a short synonym for "accounting for variation", because it is not his job to look into sociopolitical reasons behind the regularities. But a social scientist should do so. In the example above, x_1 and x_3 are obviously collinear to an appreciable degree. Which one comes first and affects the other? Or do they both result from some underlying third factor? There are many ways to tackle the problem, some of them statistical (e.g. with lagged variables). Constructing logical quantitative models is another approach. In retrospect, my first two incursions into such models for social sciences came early on. They offer a nice starting point, before bringing what I learned from studying physics.

The Battle Equations. While herding our seven cows on an Estonian farm I had plenty of leisure time and, since WWII was going on, I pondered the following. Suppose 100 men fight 50 men to death on an open field. The two sides have equal guns and skills. How many of the larger force will remain when the smaller force is destroyed? I sensed it would be more than just 100-50=50, because while the smaller force manages one hit, the larger one will already have two hits, so that the ratio changes from 100/50=2 to 99/48, which is more than 2. Thus the advantage of the larger force keeps increasing. I could not proceed much further because I had no paper with me in the pasture, and at home there were other things to do.

Much later, while working in an industrial laboratory, I recalled the problem and easily solved it, using differential equations, only to find that F. L. Lanchester had published such equations much earlier (see Lanchester 1956). They are akin to Richardson's arms race equations discussed in Section 4.6. The basic idea is to consider what happens during a short time spell, and deduce the consequences over a longer period.

Reporting of Casualties. As I was completing high school in Marrakech, the first Moroccan uprising against French colonial rule took place in Casablanca. My friend Jacques Favreau, with family in high military circles, asked me to guess how many people actually were killed. My answer was along the following lines. "The official figure reported in the newspapers is forty. Our Moroccan servant claims about four thousand. Taking the geometric mean yields about four hundred." It turned out I was off only by 100, compared to confidential army estimates.

Why did I use the geometric mean rather than the simpler arithmetic? The latter would have amounted to slightly above 2,000, regardless of whether the official figure was 4 or 400. The discrepancy between the rumored 4,000 and the official 40 clearly showed considerable over- and underamplification, respectively, of the actual figure. Not knowing which side would distort more, my best possible guess was to assume an equal distortion on both sides, by a factor of about 10. It was a model of equal distortion.

Little did I suspect that going by the geometric mean of the extremes would become the workhorse of what I consider my main accomplishment -- putting specific numbers into Duverger's statements (1951, 1954) on the number and size of parties (Sections 5 and 6). Jacques, who later commanded an armored division, said he could not think of any other person who would try to answer his question by such means. How many such people are there, almost 50 years later, among established political scientists?

It is time to proceed to the lessons of physics. My first MA (yielding a note in *Nature*, Taagepera et al. 1961) as well as my Ph.D. (resulting in Taagepera and Williams 1966) were in experimental solid state physics. I worked in applied physics until 1970, while taking night courses in political science.

Among model building techniques (used in Taagepera and Nurmia 1961), I was struck most by the ability of differential equations to start with stipulations of local conditions and end with broad overall patterns. Thus the simple statement "rate of growth is proportional to existing size" is expressed as a very simple equation, $dS/dt=kS$. Yet it produces the much more complex exponential pattern, $S=S_0 exp(kt)$, where k is rate constant and S_0 the value at time t=0. In political science around 1970, only Lewis F. Richardson's longstanding arms race equations (with major publication in 1960) applied such an approach, with several differential equations interacting, along with the aforementioned battle equations.

Connected with the former, initial conditions play a role. For instance, they determine the value of the constant S_0 in the exponential equation. Sometimes boundary conditions play a similar role, in conjunction with assumption of continuity (or quasi-continuity, in the case of granular matter). Along with basic differential equations for electricity, the nature of container walls (some conducting, some isolating) fully determines the electric field in every location. Thus various constraints help make a general equation more specific.

Another notion is minimizing of some quantity. When traversing media that impose different speeds on light, light follows a path that minimizes the total time needed. While differential models are solved by integration, here the path is opposite: the derivative of the initial equation is calculated and required to be zero. All these calculus approaches assume continuity, but summations of discreet items are often easily changed into integrations as fair approximations.

An even broader notion is squeezing maximum information out of almost complete ignorance -- ignorance-based quantitative models, as I have called them in a more detailed exposition (Taagepera 1999b). Take the growth equation $dS/dt=kS$. If k itself is a function of S, any growth pattern is included, as long as no factors other than the existing size S enter. In utter ignorance of whether k is an increasing or decreasing function of S (at a given value of S), our best guess is not to favor either direction. This means assuming that k is constant, in the absence of contrary information. The ubiquitous and powerful exponential pattern results. It rests not on information but on dearth of it.

A specific aspect of the ignorance-based models is to determine the extremes that are conceptually possible. In the absence of any other information, there is no reason to expect the median of actual cases to be closer to either conceptual extreme. E.g., in an electoral district where 25 seats are allocated by proportional representation, at least one party must win seats, and at most 25 parties. In the absence of any other information, the best guess is the geometric mean of 1 and 25, meaning 5 -- which usually comes close.

Does this reasoning "explain" the outcome any more substantively than does correlation analysis, an approach I have characterized as empirical rather than explanatory? There is one tremendous difference -- the difference between prediction and postdiction. Once data on a large number of elections in 25-seat districts are in, statistical analysis may show that the mean number of seat-winning parties is, say, 4.76, with some standard variation. Before such data are gathered, the statistician has no inkling of the outcome. In contrast, the conceptual extremes approach enables us to make a prediction on logical grounds, before any data are in. We might prefer a model with more political substance to it, but at least we have some predictive ability. And we should be cautious not to look for exclusively political reasons for phenomena that are largely determined by broader limiting frames.

There is nothing inherently physical in all these notions. This is logic, expressed in a systematic way. Physics was merely the context in which I first learned to apply them. When shifting to the social scene I noted analogous situations. The analogies are with the mathematical notions of continuity, minimization, etc. No naive analogies between social and physical worlds are involved.

For each of the issues discussed, I'll first state the problem, then present the answer (usually in the form of an equation) and supply a data graph as a reality check. Only then, when some agreement with the real world is established, will I outline the logical model. Finally, I ask what the findings are good for. The concluding Section 7 reviews the various methodological notions involved. To see in detail how the models were built, how well they fit actual data, and what their assumptions and limitations are, one must go to the original articles. I have tried to limit the references to the strict minimum; for bibliographies of previous and related work, see articles listed.

2. Effective Number of Components

Although this is not a logical model in the strict sense, I will start here, for the following reasons. The effective number of components (parties, in particular) has been my most visible and cited contribution to political science. More important, this number is a building block that enters several of the models to be described.

It is relatively easy to construct models with equal-sized components, but it is much harder to do so when some components are large and some are tiny. In this case some equivalent constellation with equal-sized components must be devised. As an example, when seats in a 100-seat assembly are distributed as 45-29-21-5, do we essentially have a two-, three-, or four-party system? The problem is analogous when the same constellation (45-29-21-5) represents the percentages of religious or ethnic groups in a country, or the distribution of its exports among other countries, or the production shares of automobile manufacturers.

The Laakso-Taagepera effective number of components (Laakso and Taagepera 1979) is

$$\mathbf{N = (\Sigma S)^2/\Sigma(S_i^2) = 1/\Sigma(s_i^2),}$$

where S is the total size, S_i the size of the i-th component, and $s_i=S_i/S$ its fractional share. The example above is chosen so that N=3.00 exactly, suggesting that the constellation 45-29-21-5 is somehow equivalent to 33.3-33.3-33.3. Most often N is a non-integer, making such direct interpretation fuzzier.

There are other ways to express an effective number of components (see Taagepera and Shugart 1989: 259-260 and Taagepera 1999c), but N has become the standard measure in political science (Lijphart 1994: 70; Cox 1997: 29). It is not utterly new, having analogues in physics. In economics the "equivalent number of firms" has been mentioned at least since the 1950s, but this notion has offered less interest to economists than the number of parties has for political scientists. Economists and others have been using for a long time the Herfindahl-Hirschman concentration index, which is the inverse of *N*: $HH=S(s_i^2)=1/N$. Hence it carries exactly the same information content. So does the Rae (1967) fractionalization index $F=1-HH=1-1/N$.

Why use *N* rather than *HH*, which is simpler to calculate? It's that *N* has more intuitive content. N=3.2 evokes the image of approximately three parties, while the corresponding HH=0.31 or F=0.69 evoke no such images. This becomes important in building logical models. It will be seen that various models can be built on the basis of the number of parties, while I have noticed no ways to start from *HH* or *F*. The latter are abstract indices, while *N* can be visualized.

Of course, *N* has its limitations.[3] When one looks at coalition potential (assuming a need for absolute majority), other approaches are advisable. In particular, *N* can be misleading in cases like 53-17-9-9-8-3-1 that yields N=3.00, intimating three fairly equal parties when actually one party has absolute majority. In such cases *N* should be supplemented by the inverse of the largest share (1/0.53=1.89), which is another measure of the number of parties. Its advantage is that it falls below 2 whenever one party predominates (Taagepera 1999c), thus automatically signaling an absolute majority. But under other circumstances *N* gives us more information than the inverse of the largest share can.[4]

3. The Effects of Country Size

The importance of country size was stressed early on by Robert Dahl and Edward Tufte in their *Size and Democracy* (1973). In particular, they noted empirical connections with parliamentary size (1973: 80-84), foreign trade as percentage of GNP (1973: 113-116), and military expenditures as percentage of GNP (1973: 122-128). For assembly size, a logical quantitative model had by this time already been formulated and tested (Taagepera 1972). A model for trade came later (Taagepera 1976a). The size dependence of military expenditures remains to be explained, as far as I know. So do other issues, such as what determines the number of administrative subunits in a country. On the other hand, a model for sizes of cities in a given country (Taagepera and Kaskla 2001) extends, surprisingly but logically, to citation shares of political science journals (Taagepera 2001c,d).

3.1. The Cube Root Law of Assembly Sizes (first or only chambers)

The number of seats (*F*) in first or only chambers of democratic legislative assemblies depends on population represented (*P*) as

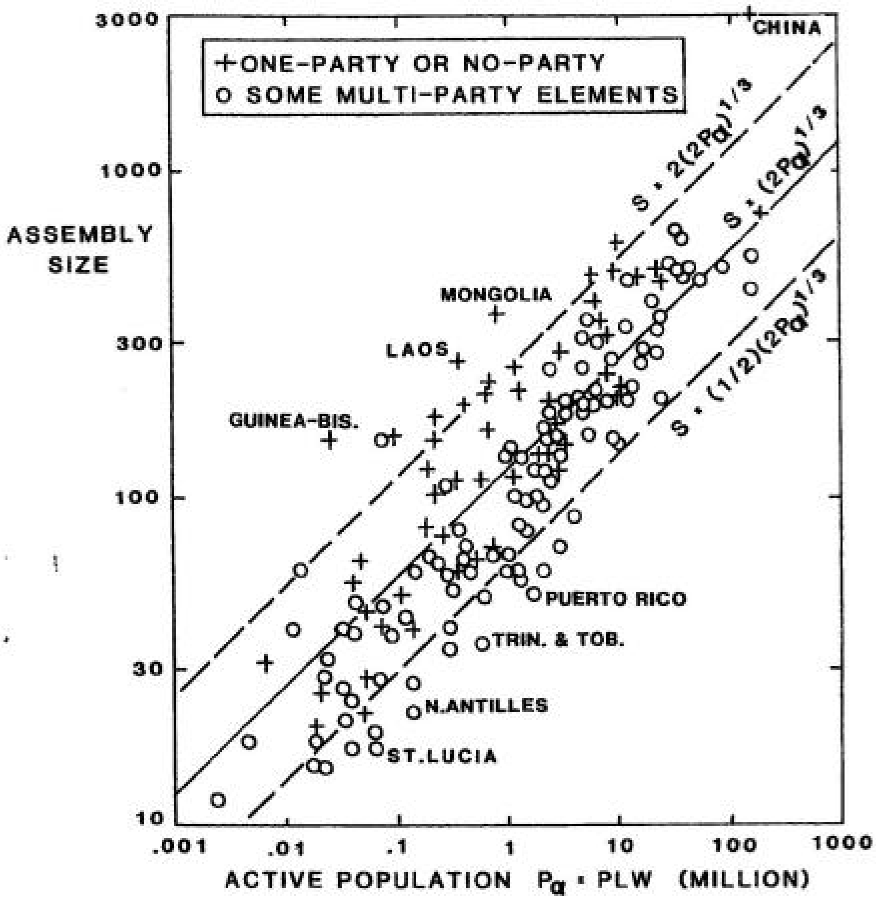
$$\mathbf{F=P^{1/3}},$$

as a first approximation (Taagepera 1972). E.g., for P=8 million, one would expect F=200. When the population is unusually young and largely illiterate, assemblies tend to be smaller. Hence a second approximation for the cube root law is

$$\mathbf{F=(2LWP)^{1/3}},$$

where *W* is the working-age share of the population and *L* is the fraction of adult population that is literate (Taagepera 1972). In this form, the law applies within a factor of 2 to most countries with populations over 500,000 (see Figure 1). Even with the second approximation, countries of less than 500,000 still tend to have assemblies smaller than expected.[5] Communist one-party states tend to have larger assemblies than expected. The overall correlation between log*F* and log*P* is around R^2=0.8 (Taagepera and Recchia 2002).

FIGURE 1. Assembly Size vs. Active (adult literate) Population



Source: From Taagepera and Shugart (1989: 178).

The logical quantitative model behind the cube root law is based on minimization of the total number of three types of communication channels: from representative to constituent, from one representatives to another, and monitoring interactions between two other representatives. A basic building stone that will be encountered again in other contexts is the number of communication channels (c) between p actors. It is easy to show that it is $c=p(p-1)/2$, which for large p approximates to $c=p^2/2$. Given that communication is what turns isolated individuals into society, this equation must be among the most fundamental in social sciences.[6]

Does the number of seats in first or only chambers matter in politics? The number of parties represented and the degree of proportionality of representation is demonstrably reduced in assemblies of less than 100 members -- meaning most countries below 1 million (Lijphart 1994). Further ways in which assembly size may matter are offered in Dahl and Tufte (1973). Larger size can be expected to make communication more cumbersome, but to an extent that is hard to pin down. We'll see that assembly size becomes a significant building block when constructing a model of representation.

3.2. The Size of Second Chambers

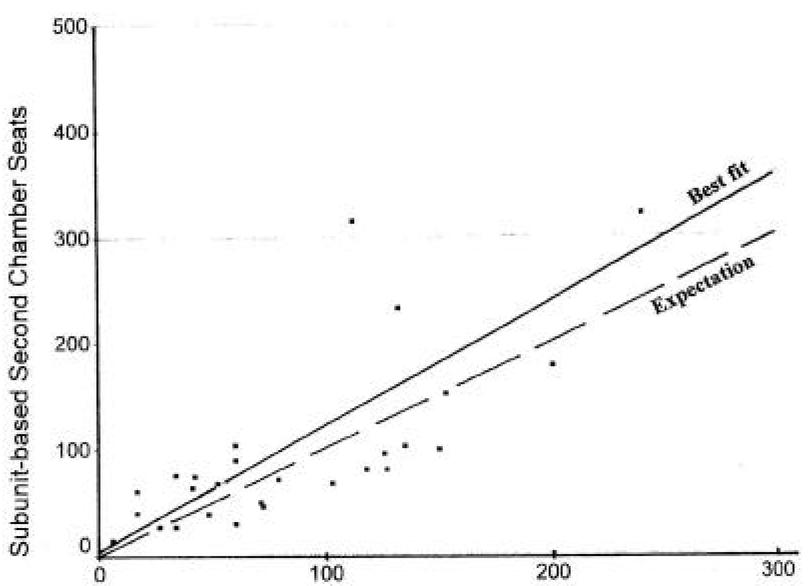
Second chambers are constituted on quite varying grounds. They tend to have fewer seats than the corresponding first chambers, but no general model has been devised. The best empirical fit of second chamber sizes (S) to first chamber sizes (F) is $S=1.00F^{0.79}$, with R^2=0.68 (Taagepera and Recchia 2002). The connection to total population (P) seems more remote: $S=0.46P^{0.30}$, with R^2=0.55.

A logical model can be constructed for a subset of second chambers -- those where the seats are based on territorial subunits (federal or other). If the number of such subunits is *n* and the first chamber has *F* seats, then the best guess for second chamber seats, in the absence of any other knowledge, is

$$\mathbf{S=(nF)^{0.5}}$$

(Taagepera and Recchia 2002). The empirical best fit between *S* and $(nF)^{0.5}$ is $S=1.18(nF)^{0.5}+6$, with $R^2=0.45$ (see Figure 2).

FIGURE 2. Number of Subunit-based Second Chamber Seats vs. the Geometric Mean of First Size Chamber Size and the Number of Subunits.



Source: From Taagepera and Recchia (2002).

This equation results from an ignorance-based model based on two boundary conditions. 1) To represent *n* units, at least *n* seats are needed -- this lower limit would represent pure representation of territorial units regardless of their populations. 2) A size equal to that of the first chamber (*S=F*) is taken as the upper limit -- this would be pure representation of the total population. Taking equally into account both units and their populations leads to the geometric mean of *n* and *F*. Both conceptual extremes (*S=n* and *S=F*) do occur, but $(nF)^{0.5}$ is close the actual world median.[7] This model also raises a further question: What determines the number of territorial subunits in a country? I have no answer, as yet.

Given the weakness of most second chambers, their number of seats may matter in politics even less than is the case for the first chambers. However, the limited model presented here may offer some guidelines when new quasi-federal structures are set up. One such timely issue is presented by the assemblies in the European Union.

3.3. The Size of European Assemblies

The European Parliament (EP) is organized by party groupings that transcend countries. Thus it arguably represents the population of European Union (EU) as individuals and thus is akin to a first chamber. Indeed, EP size has edged ever closer to the cube root of the EU population. The enlargement expected in 2003 would raise EP to 732 seats, while the cube root of the corresponding population (481 million) is 784 (Taagepera and Recchia 2002).

In contrast, the Council of the EU (CEU) explicitly represents people through their countries and hence is arguably akin to a federal second chamber that represents the countries, yet allocates more votes to larger countries. Indeed, from 1979 to 1999, the total number of votes in the CEU was within 10% of the value predicted by the model $S=(nF)^{0.5}$, if S is the number of CEU votes, n is the number of countries and F is the number of EP seats (Taagepera and Recchia 2002). However, the Treaty of Nice in 2000 boosted the CEU to a size comparable to that of EP, which may not be functional. The usefulness of the logical models for first and second chamber sizes lies in pointing out such discrepancies.

3.4. Trade/GNP Ratio

This is a quite different effect of country size, with relevance for political economy. Smaller countries tend to have larger Import/GNP and Export/GNP ratios than larger countries. The reason becomes clear when one considers conceptual extreme cases. A country encompassing the entire world would by definition have zero external trade, while a country of one person would have Imp/GNP and Exp/GNP ratios of 1.00 (and more in the case of transit trade). Assuming continuous change in-between these limits, trade/GNP must decrease with increasing country size. The question is: How fast does it decrease?

Empirically, the following simplified equation was found to apply within a factor of 2 in 94% of the cases (R^2=0.69):

$$\mathbf{Imports/GNP=40/P^{1/3}.}$$

where P is country population (Taagepera and Hayes 1977). Conveniently enough, if P is in million, the Imp/GNP ratio comes in per cent. (This would not be the case if the exponent were different from 1/3.) The occurrence of the cube root of population is reminiscent of the cube root law of assembly sizes, but this is coincidental. For exports, an equation close to $Imp/GNP=30/P^{1/3}$ applies to a poorer degree: only 77% of the countries are within a factor of 2. The shortfall of exports compared to imports is made up by tourism, export of labor, shipping, banking services, foreign aid, etc. The simplified combined equation

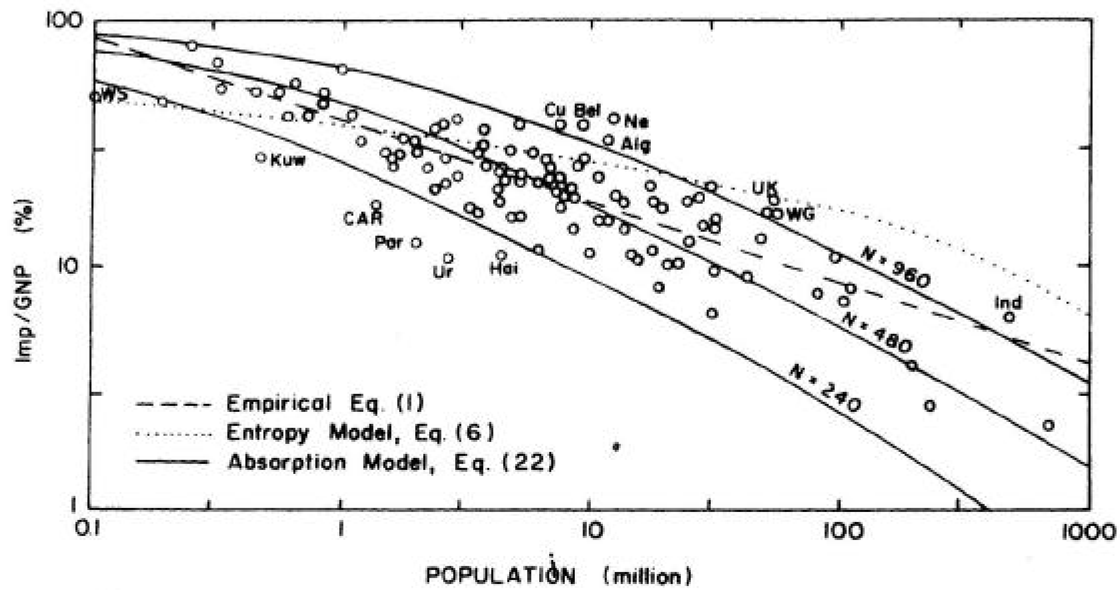
$$\mathbf{Trade/GNP=70/P^{1/3}}$$

applies within a factor of 2 in 92% of the cases (Taagepera 1976a; Taagepera and Hayes 1977). The pattern is broadly confirmed by several other studies. In particular, the best fit by Dahl and Tufte (1973: 114) corresponds to $Trade/GNP=66/P^{0.29}$ where P must be in millions and the ratio in per cent. Dahl and Tufte argue convincingly why the share of foreign trade should decrease with increasing population. But why should it decrease with an exponent around 0.30 rather than, say, 0.10 or 0.50? We need a logical quantitative model.

The model proposed (Taagepera 1976a) proceeds from the one used in physics for absorption of any flux of particles in homogeneous surroundings. The relative change in flow intensity (I) with distance (r) is constant: $dI/dr=-kI$, where k is the absorption constant. In one-dimensional space this differential equation integrates into an exponential function: $I=I_0 exp(-kr)$, where I_0 is the intensity at the source. Many simplifying assumptions enter. I assume that goods produced flow out of such sources (factories etc.) and are absorbed by customers. Whatever flow continues beyond a country border is called export. It is a matter of integrating the production and absorption of many sources, assumed to spread uniformly across the country.

In one dimension the problem is quite simple, but for two-dimensional countries only an approximate solution is possible, as given in Taagepera (1976a). The central constant, akin to constant k above, is the "characteristic absorption number". It is the number of people that a flow of goods encounters before its intensity is reduced to 1/e=0.37 of its original intensity. The best fit to data, fairly close to the empirical $Imports/GNP=40/P^{1/3}$ in the population range of actual countries, is found with population number around 500 (see Figure 3).[8] For countries of more than 100 million the model falls below the simple empirical fit, in agreement with the data points for US, USSR and China, but not India.[9]

FIGURE 3. Import/GNP Ration versus Population, on log-log scale



Source: From Taagepera (1976a).

What is the model good for? It helps to evaluate whether a country's trade is out of line given its size. At face value, China's low Trade/GNP ratio (around 1970) suggests isolation, but in reality it merely expresses the fact that China has a large population. Thus the model helps us compare realistically the trade activities of countries of disparate sizes -- and also the world at different time periods. If country-to-country trade within the European Union comes to be considered domestic trade, then the world total trade figures will drop drastically, unless corrected using a model like the present one.

3.5. The City-country Rule

In all previous models in this section country population was a given starting point. In the following, in contrast, it is constructed by adding up the populations of cities and other settlements. Once more we shift away from politics, this time in the direction of human geography.

The starting point is the well-known rank-size rule, often called Zipf's law.[10] It has long been observed that the population (P_R) of the R-th ranking city in a country tends to be around $P_R=P_1/R$, where P_1 is the population of the largest city. More complex variations exist, but let us focus on the simple expression. It uses one extreme anchor point -- the largest settlement. However, let us also consider the other conceptual extreme -- the smallest conceivable settlement, an isolated forest hut inhabited by a single person.

The model proposed (Taagepera and Kaskla 2001) assumes that the rank-size rule applies all the way down to this hypothetical one-person settlement, thus accounting for the entire population of the country. What is the total population (P) that would result? The answer is

$$\mathbf{P= P_1 ln P_1.}$$

This "city-country rule" predicts that the total population of a country is the largest city's population multiplied by its natural logarithm.

How is this city-country rule is obtained? One sums up the populations of all settlements, from rank 1 to the rank of the smallest settlement of one person. The previous rank-size rule $P_R=P_1/R$ implies that the **rank** of that smallest settlement equals the **population** of the largest city (P_1). The summation is then approximated by integration of P_1/R, and $P = P_1 \ln P_1$ results.

Since the largest city often deviates from the rank-size rule (the so-called primacy factor), one often obtains a better fit with the second- or third-largest city, using the more general expression

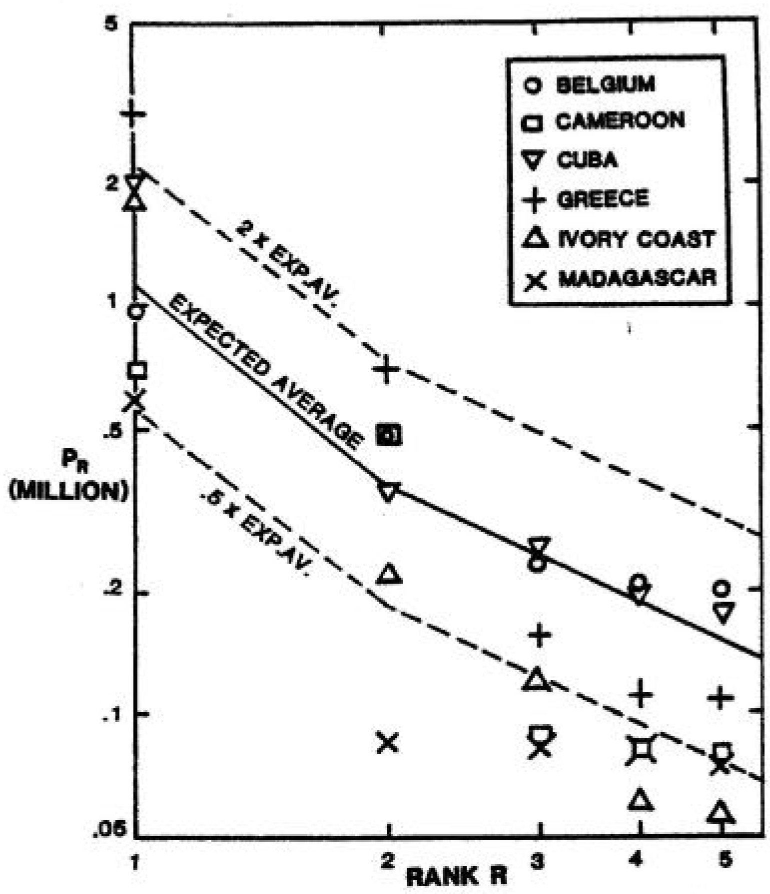
$$\mathbf{P = RP_R \ln(RP_R),}$$

where P is the population of the country and P_R is the population of the R-th ranked city.

The model fits well, on the average, for countries of less than 100 million. For larger countries a more elaborate formula holds. The largest city in the country, subject to the aforementioned primacy factor, needs a correction factor that vanishes for large countries.

The important result is that we can use the equation in the reverse direction and predict the populations of all cities in a country, once the total population of the country is given. We can quickly tell whether the actual city populations are large or small, compared to the world average in countries of comparable size. Figure 4 shows the six countries with populations closest to exactly 10 million (Taagepera and Kaskla 2001).[11] If the deviations from expectation look large, please keep in mind that this is the only model proposed, as far as I know, to deduce city sizes from the total population of a country. We are making progress.

FIGURE 4. City Populations in Six Countries with Exactly 10 Million Population: Actual and as Predicted by City-County Rule with Primacy Adjustment. Both Axes Use Logarithmic Scales



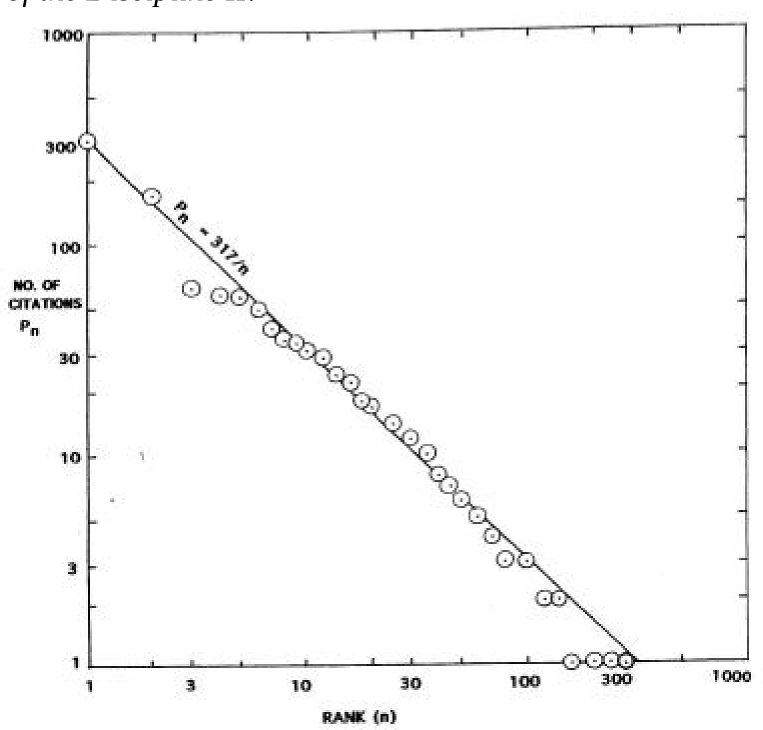
Source: From Taagepera and Kaskla (2001)

3.6. From Cities to Journals

The city-country rule applies in principle to other systems of defined total size and defined smallest conceivable component. One example is the number of citations to various journals.

The chapter bibliographies in *Political Science: The State of the Discipline II* (Finifter 1993) have a total of 1834 citations to journals (excluding books and other non-periodicals). On that basis the city-country rule predicts 318 different journals cited at least once, and also 318 citations for the topmost journal. The actual figures are 281 periodicals cited, and 317 citations for *APSR* (Taagepera, 2001c). Figure 5 shows that agreement is good at all ranks, except for a shortfall for the 4th- and 5th-ranking journals. The fit is almost as good for another overview, *A New Handbook of Political Science* (Goodin and Klingemann 1996).

FIGURE 5. Hyperbolic Relation between the Frequency of Citations and Ranking of Journals, Based on the *State of the Discipline II*.



Source: From Taagepera (2001c).

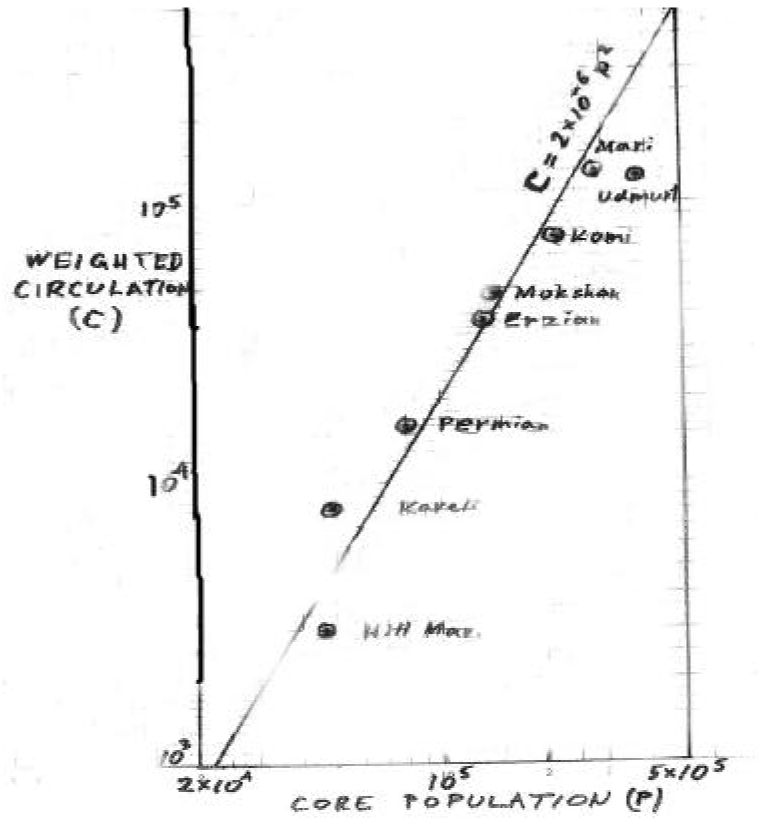
In contrast to geographical settlements, here the rule can be tested literally down to the smallest conceivable component.[12] But the basic question remains: What is the reason for such empirical regularity as the rank-size rule? Why should it hold? Once more, the answer comes from an ignorance-based model (Taagepera 2001d). At the moment, the assumptions made in the model are satisfied by journal articles but not quite so by small settlements. It's a beginning of an explanation.

We have shifted far away from the effects of country size -- and in a way closer to home for political scientists. How does the regularity observed (and now partly explained) matter, outside the city-country context? Directly, it remains a mere curiosity. But curiosity is the driving wheel of science. Here we have an example of how a model devised for one purpose applies elsewhere--which means that we have a phenomenon of considerable generality. This generality reinforces the need for understanding the logical reasons behind the rank-size rule.

3.7. The Number of Speakers and the Cultural Vigor of Languages

In a world where English is rapidly becoming a universal *lingua franca*, what is the outlook for other languages? A language spoken and/or read by 100 million is clearly likely to preserve a wider scope than one used by 10 million. And with one million speakers a language has better survival chances than with 100,000. But by how much?

FIGURE 6. Weighted Circulation vs. Core Population of Finno-Ugric Peoples in the Russian Federation.



Source: From Taagepera (1999a: 401).

To be more specific, consider the total yearly circulation of periodicals published in a language (*C*). With the population (*P*) doubled, the existing journals double their potential market, but this is not all. New, more specialized periodicals can be published, instead of people having to read on such topics in another language. Hence one would expect $C=kP^n$, with $n>1$ and k a constant. The number of communication channels between *P* people is close to *P* squared (see Section 3.1), and periodicals are a form of communication. Hence one would expect

$$\mathbf{C=kP^2.}$$

I have been able to test it with only one data set (Taagepera 1999a), because in general the condition "all other things being the same" is hard to satisfy. The model $C=kP^2$ fits rather well, with $k=2x10^{-6}$ (see Figure 6).[13]

4. Changes over Time: Growth, Decline and Duration

In this section we proceed from world population to hyperinflation, the number of polities in the world, and arms races. The common feature is interactive change over time. Growth of Western civilization and the fading of older periods of history in our minds are simpler processes in principle but harder to pin down in practice. Duration of cabinet coalitions follows a different logic.

4.1. World Population Growth

The basic growth pattern of self-sustaining systems is exponential, meaning constant relative (percent) growth. As mentioned earlier, it fits an ignorance-based model: In the absence of any other knowledge on whether the growth rate is increasing or decreasing, the safest median bet is that it remains constant. Most growth processes slow down after a while, because they encounter food and space limitations or internal checks. Such input represents addition of further information to be built into the model.

The exponential equation $dP/dt=kP$ means growth without limits. The so-called simple logistic equation offers the simplest way to model growth with limits, by making the previous constant k itself decrease linearly with increasing size: $dP/dt=k(1-P/M)P$, where M is the maximum sustainable size. The human population, however, has increased with an *increasing* growth "constant" k for the last thousand and most likely a million years, reversing gears only a few decades ago. Hyperinflation is the only other phenomenon I know that has an increasing growth constant.

The simple model to express super-exponential population growth consists of two exponentials.[14] Basically, population still grows at a rate proportional to the existing population: $dP/dt=kP$. But at the same time, technology grows at a rate proportional to the existing level of technology: $dT/dt=hT$. Moreover, the two exponentials interact. A higher level of technology boosts the population growth "constant" (k), because the same area can now support more people. And a larger population of potential inventors boosts the technology growth "constant" (h). Thus both population and technology grow faster than exponentially.

In mathematical terms (Taagepera 1976b), population grows as $dP/dt=k[T]P$, and technology grows as $dT/dt=h[P]T$. Assuming that the functions $k[T]$ and $h[P]$ are positive power functions and that zero population means zero technology, integration yields a quasi-hyperbolic equation:

$$\mathbf{P=A/(D-t)^M},$$

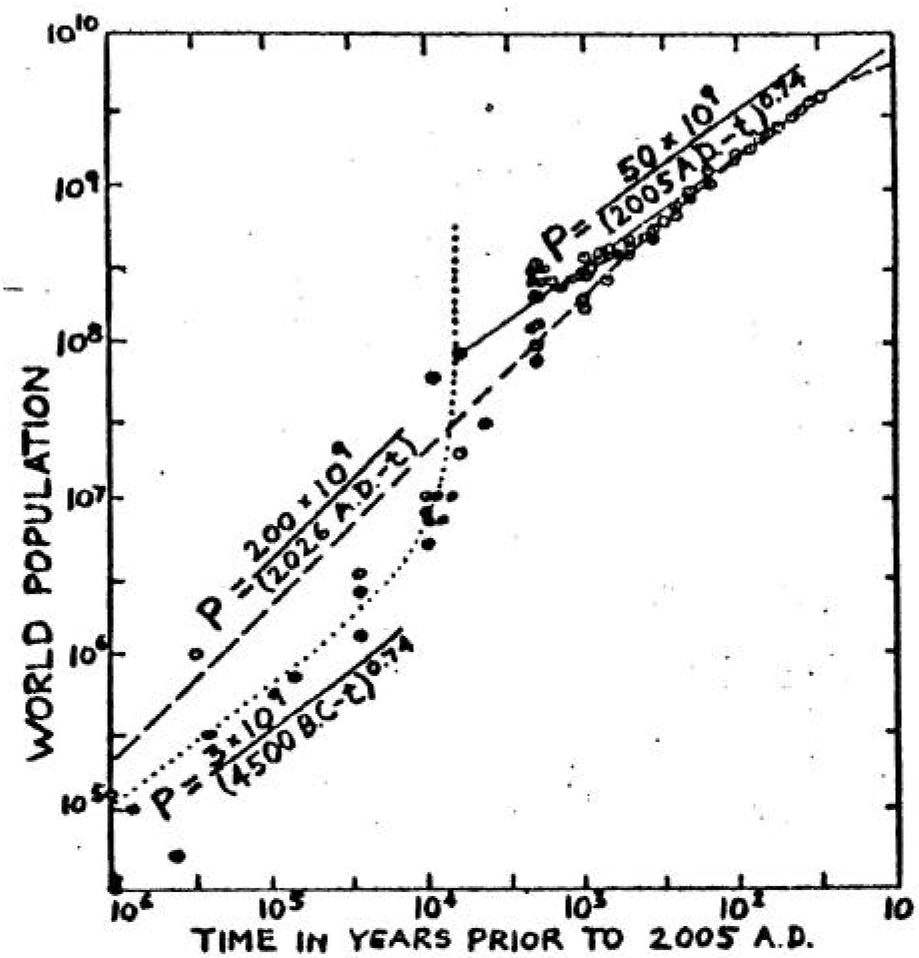
where A, D and M are constants, and a similar equation for technology, which is harder to test. For population,

$$\mathbf{P=50\ billion/(2005-t)^{0.74}}$$

agreed with various estimates within 30%, from -4000 to +1970. It is a quasi-hyperbolic pattern where population would shoot up to infinity at "doomsday time" $t=D$, unless something changes. From the vantage point of 1976, D=2005 was disquietingly close: Something had to change very suddenly.

Several other authors (Meyer 1974, in particular; cf. Taagepera 1979b) observed the same pattern on empirical grounds, but with doomsday around 2026. Indeed, such a constant yields a better fit for population estimates as far back as one million years. On closer look (Taagepera 1979b), I noticed a discontinuity during the Neolithic agricultural revolution. Up to -4500, the best fit is with $P=3\ billion/(-4500-t)^{0.74}$, intimating some sort of a major shift around -4500-- see Figure 7.

FIGURE 7. World Population (log/log scale)



Source: From Taagepera (1979b)

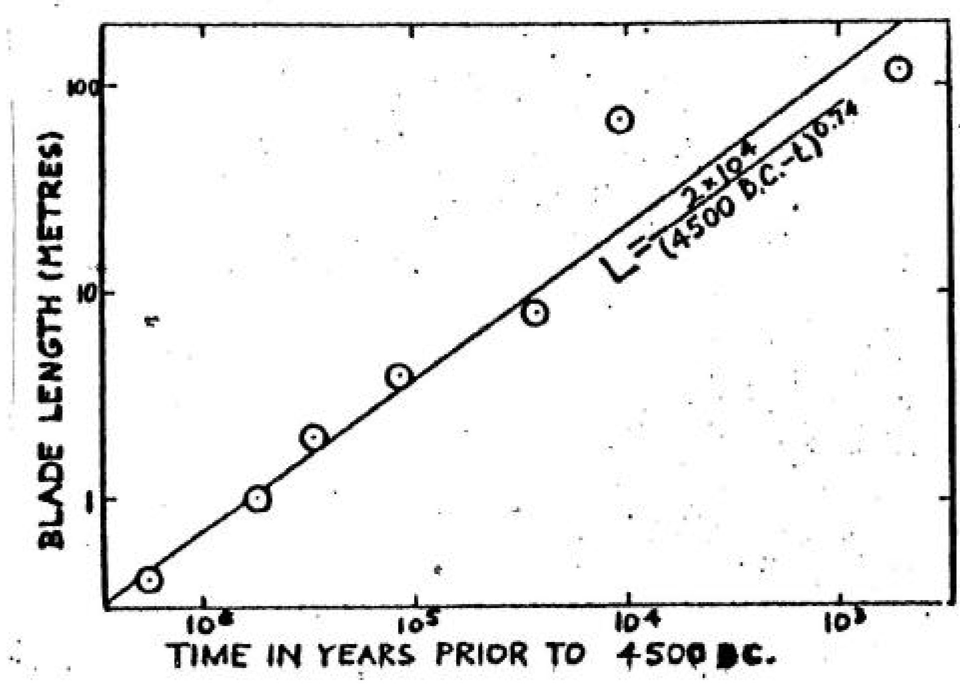
For the same time period a simple measure of technology is available: the blade length obtained from 1 kg of silex. Up to -5300, this length, as estimated by archeologists (see Meyer 1974: 39-40), fits well a quasi-hyperbolic equation using the same constants *D* and *M* as the corresponding population equation: $T=20{,}000\ meters/(-4500-t)^{0.74}$ -- see Figure 8. What happened around this earlier "doomsday " of -4500? Stone began to be replaced by metal, but also several early civilizations went under. This *"Neolithic singularity"* deserves further study.

Later, I devised a simple graphical way to determine the doomsday date (*D*) and visualize the momentous issue involved. The quasi-hyperbolic patterns become straight lines when graphed log*P* vs. log*(D-t)*, but this presupposes finding first the proper value of *D* by trial and error. This inelegant approach is circumvented by shifting to population doubling time (T_2) after a given date. The quasi-hyperbolic model predicts a linear decrease in T_2, hitting $T_2=0$ at time *D*. Within the variation range of various population estimates, the line

$$\mathbf{T_2=0.53(2015-t)}$$

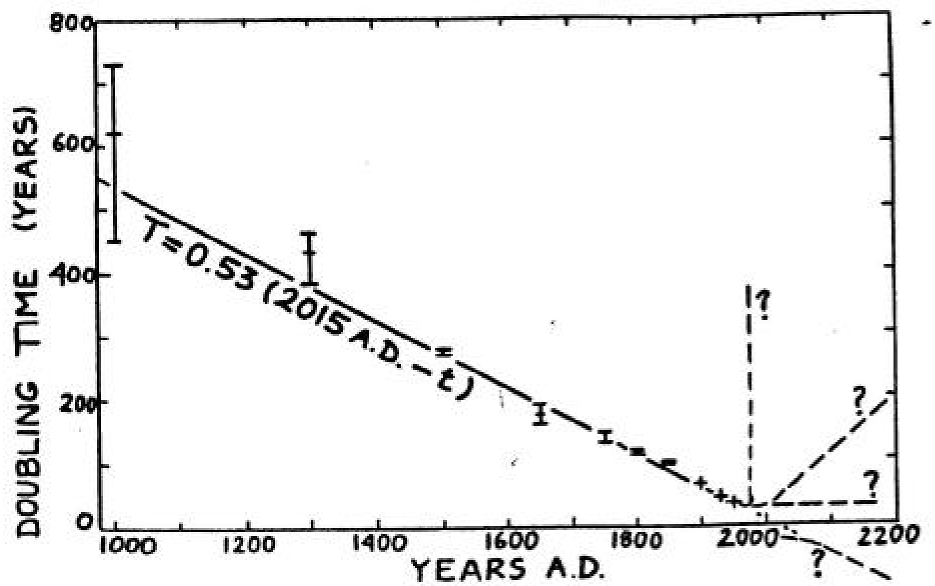
fits for the last 1000 years, up to 1950 -- see Figure 9. Extending this best-fit line back to the most distant estimates available (900,000 years) still yields a marginally satisfactory agreement (Taagepera 1981).[15]

FIGURE 8. Blade Length Obtained from 1kg of Silex (log/log scale)



Source: Taagepera (1979b)

FIGURE 9. The Strange Case of the Vanishing Doubling Time of World Population



Source: From Taagepera (1981).

Populations cannot grow to infinite sizes because of limited resources (including space). Thus I reworked the simple interaction of population and technology, adding the rate of depletion of resources (dR/dt), which increases with larger populations and may or may not increase with higher technology. The resulting equations, too complex to be reproduced here, are shown in Taagepera (1979b). Resource depletion eventually imposes a decrease in population growth rate (dP/dt), as has been the case since about 1975. *A million-year trend has reached an end.* Testing these more complex equations can be carried out meaningfully only when population figures far beyond the turning point of 1975 become available, but a start could be made now.

Remember that the simple interaction model assumed that more people meant more potential inventors. This may not hold when people become so numerous that many inventors merely duplicate each other's findings. In this case the model suggests a shift from quasi-exponential (with a doomsday date) to "doubly exponential", still super-exponential but with no finite doomsday date. Population growth since 1900, when world population reached 1 billion, is compatible with such a simplification (Taagepera 1979b).

How do these models matter? They highlight the potentially catastrophic nature of world population explosion, which still continues. But this is well visible on empirical grounds too. What the models add is explanation of the phenomenon in terms of interaction of population, technology, and resource squeeze. Such explanation may help us to alleviate what I consider the greatest danger facing humankind in centuries to come: a catastrophic population collapse resulting from overpopulation, depletion of resources (including clean air) and warfare over the last remaining resources. The problem is in the realm of demography, but solutions involve politics. Meanwhile, increasing populations affect governability in cities and entire countries.

4.2. Hyperinflation

As mentioned, hyperinflation is the only other phenomenon I know that shows a super-exponential pattern. Here the interaction is between loss of value by currency and loss of trust in currency by people. It looks simple enough, but somehow I could not formulate the suitable differential equations. Drafting the present overview made me attempt it once more, and I got it! -- only to find that the resulting equation is essentially equivalent Cagan's (1956) money demand function. Just in case my approach might cast some additional light on the issue, it is briefly presented here.

The price index (*p*) is what is tabulated in data sources, reasonably enough. But whether the absolute prices are in thousands of yens or in tens of dollars does not matter, as long as *p* does not change abruptly. What people do notice and react to is the relative (percent) rate of change of prices, $r=(dp/dt)/p$. The faster prices go up (meaning high *r*), the faster people try to get rid of money on hand, and this increasing turnover boosts *r*. Hence $dr/dt=kr$, so that *r* increases exponentially ($r=ae^{kt}$), and prices themselves increase as

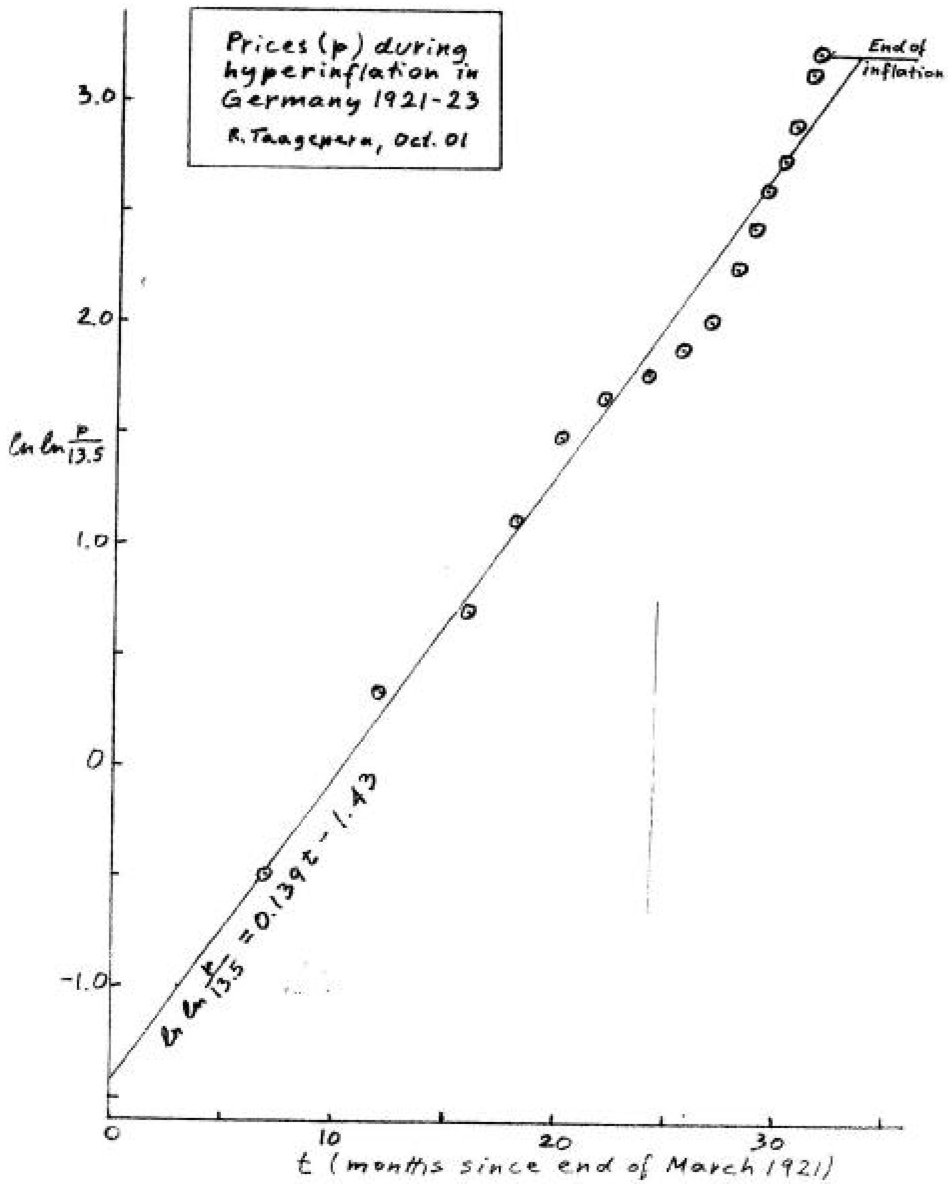
$$\mathbf{p=b\ exp(ae^{kt}).}$$

The growth in prices is steeper than exponential (decreasing doubling times) but without shooting to infinity at a final time (like the quasi-hyperbolic model). This is the same "doubly exponential" model that emerges in population growth, when increasing population no longer affects the number of non-duplicating inventors.

The model is best tested by taking logarithms twice. With the constant *b* properly chosen, this leads to a linear relationship: $ln\ ln(p/b)=kt-A$, where $A=e^a$. Figure 10 shows the degree of fit for the famous German hyperinflation of 1921-1923. The fit is even better for the preceding German inflation of 1919-1920, which looks mild in comparison but actually annihilated the prewar savings, reducing their value to 1/17 of the original.

What is this model good for? I have no overview of what economists have done with it. What seems of interest to me is that it would enable us to compare the rate constants *k* for various cases of hyperinflation. For Germany, k=0.25 per month during the relatively mild inflation of 1919-20. Surprisingly, it was less during the infamous hyper-inflation of 1921-23: k=0.14 per month.

FIGURE 10. Prices during Hyperinflation in Germany, 1921-23



4.3. Size and Effective Number of Polities

Over 5000 years of recorded history, the areas of the largest political entities have tended to increase. Accordingly, the effective number (see Section 2) of separate polities has decreased. We can measure that it has decreased from close to a million to 24, when going by geographical area. When going by population, the effective number of polities has decreased from about a thousand to 15 (Taagepera 1997b). What has caused this concentration of politics? Compared to the millennial trend, are the present large polities overly large or overly small? If this trend continues, when can we expect a single world government to materialize?

Consider the effective number of polities both in terms of area (N_A) and population (N_P). We may expect that N_P tends to be the square root of N_A --

$$\mathbf{N_P=N_A^{0.5},}$$

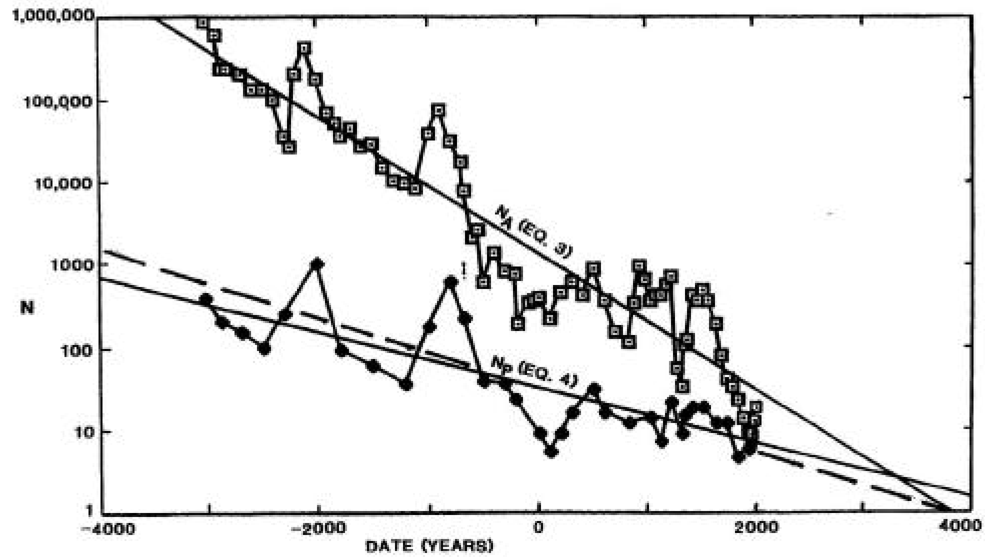
for the following reasons (Taagepera 1997b). Large polities tend to include the densest population areas. Hence N_P is smaller than N_A, which is the upper limit on N_P. The conceivable lower limit on N_P approaches 1. This would be the case when a single polity in one fertile river valley (a "super-Nile") encloses most of humankind. The Egyptian Old Empire came close to this situation 4700 years ago. In the absence of any other knowledge, we should expect N_P to be the geometric mean of N_A and 1, meaning the square root of N_A. Obviously, when $N_A=1$ we must also have $N_P=1$. The square root relationship satisfies this condition.

In the absence of any further knowledge, the best guess for the pattern of decrease of N over time is that its rate of decrease is proportional to N itself: $dN/dt=-kN$. This leads to exponential decrease:

$$\mathbf{N_A=a\ e^{-kt}} \quad \text{and} \quad \mathbf{N_P=a^{0.5}e^{-kt/2}},$$

where the constants a and k are to be determined empirically.

FIGURE 11. Effective Number of Polities, Based on Area and on Population



Source: From Taagepera (1997b) with dashed line at square root of N_A added.

This model of two interlocked exponentials works indeed -- see Figure 11. The overall pattern of $\log N_A$ vs. time is linear, corresponding to an exponential decline

$$\mathbf{N_A=1300exp(-0.19t)} \qquad \text{["EQ. 3" in Figure 11]}$$

when t is time in **centuries**. Here $R^2=0.90$. The square root model would then predict $N_p=36exp(-0.095t)$ -- the dashed line in Figure 11. This is very close to the actual best fit for N_P, which is

$$\mathbf{N_P=31exp(-0.08t),} \qquad \text{["EQ. 4" in Figure 11]}$$

with $R^2=0.68$. Population estimates of polities are second order estimates (compared to areas) and hence more error is to be expected.[16] What determines these rates of 19% and 8% per century? Why should the effective number of polities tend to decrease at such average rate, and not faster or slower? It may have something to do with increase in information flux. I have no answer, yet.

The modeling actually was preceded by considerable empirical work. The area growth and decline curves of the largest polities at all time periods had to be measured and graphed, and only one brief previous attempt (Hart 1945) existed. This kept me busy for quite a while (Taagepera 1968, 1978a,b, 1979c, 1997b). Only then could the effective numbers be measured and graphed against time (Taagepera 1997b).[17]

What is the use of such mapping of secular trends? For one, it offers a cautious answer to the question of when one might expect the advent of a single world state. According to the equation above, $N_A=1$ will be reached in +3800 (when the best fit line for N_P projects to $N_P=1.5$). Within the range of fluctuations observed, a single word state has a nil probability prior to +2600. These are more remote dates than suggested by most other researchers (see references in Taagepera 1997b). A completely bipolar world ($N_A=2$) could briefly occur by +2200 but has a 50-50 chance only by +3400. Of course, all extrapolations are speculative as long as empirical measurements are not supported by a logical explanation for the value of rate constant *k*. Meanwhile, when guessing at when "history will end", the best we can do is look at all the history we've got and extrapolate, very skeptically.

Another use of mapping secular trends is to guess at the prospects of present large polities. Compared to the average trend, we are emerging from one of the fluctuations toward low *N*, meaning unusually large empires (British, in particular) -- cf. Figure 11. The world still is slightly over-concentrated compared to the average secular trend, and earlier history has repeatedly seen a pendulum movement from over-concentration to over-fragmentation. This bodes ill for some of the largest present polities, the largest being Russia.

4.4. Will Russia Last?

Before addressing that issue, potential regularities in duration of polities must be considered. There is no evidence that durability is tied to size. However, rapid rise may lead to shoddy structure and hence early demise. Rise (*R*) and duration (*D*) times for individual polities can be defined: *R* as the time span to rise from 20 % to 80 % of eventual maximum, and *D* as the time span at one-half the maximum size. Russia has by far the longest rise time in history (235 years, after 1555), followed by Han (160 years) and several empires at 140 years. The median duration has been 160 years, and only 16 polities have lasted 300 years or more. The record (D=700 years) belongs to the Parthian-Sassanid continuum. Russia already has lasted for 335 years at more than one-half its maximum area (starting in 1665). The Soviet debacle left Russia with one-half of the Soviet population but still 75% of the area of the USSR (and the late tsarist empire).

My early work suggested that $D=3R$, and I even had a model of sorts to support that. But the more polities I measured, using more refined operationalization, the more the pattern became blurred. The best empirical fit in Taagepera (1997b) is around $D=25R^{0.5}$, where *R* and *D* are in years, but the correlation is extremely weak.

Having already lasted over 300 years, Russia has entered the perilous zone, but its record rise time improves its chances. The empirical average of $D=25R^{0.5}$ would offer Russia a duration up to year 2050. Meanwhile, however, it could continue to contract appreciably at the margins and still remain above one-half its former maximum area. A new increase in area, though possible, is less likely. Polities that lose momentum rarely recover it. For implications of a breakup of Russia, see Taagepera (2000, 2001a).

4.5. Arms Races

This is the first and indeed the only marked case where differential equation models were used in political science more than 40 years ago. Lewis F. Richardson formulated his famous equations in the aftermath of WWI, but his work was widely noticed only through the posthumous *Arms and Insecurity* (1960). I approached it enthusiastically. Unfortunately, I located some intractable problems with the model, which made me give up this line of research after 1980.

Richardson's basic model for the arms budgets (*x*, *y*) of two mutually distrustful countries is that increases in budgets are encouraged by the other side's budget but held back by one's own budget (which is a burden on economy):

$dx/dt=ky-ax+g$
$dy/dt=lx-by+h$,

where *k*, *l*, *a* and *b* are positive constants, while the inherent trust/distrust constants *g* and *h* can be negative or positive. Richardson (1960:32) solved these equations only partly, first assuming $k=l$ and $a=b$ (which is fair enough) and then adding the two equations. With $S=x+y$, the result is $dS/dt=(k-a)S+(g+h)$, leading to a simple exponential solution. And Richardson found an excellent fit for 1909-1913.

Two serious problems were pointed out in an article I published with two undergraduate students (Wagner, Perkins and Taagepera, 1975). Summations tend to smooth out the data, making the fit look better than it is. A more serious objection is that exactly the same equation for *S* would result if the two countries ignored each other and based their budget changes only on their own previous budget -- the usual internal bureaucratic pressure. Indeed,

$dx/dt=(k-a)x+g$
$dy/dt=(k-a)y+h$

also lead to $dS/dt=(k-a)S+(g+h)$. Mutual stimulation could not be distinguished from self-stimulation!

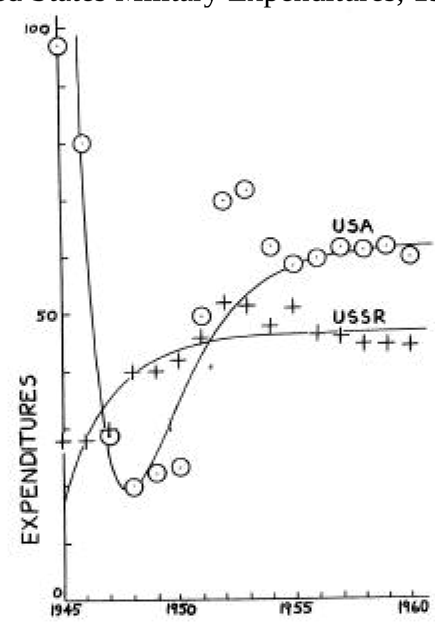
Instead of the sum, one could also use the difference $D=x-y$. The result is $dD/dt=-(k+a)D+(g-h)$, and now the 1909-1913 data are scattered all over the place (see graph in Wagner, Perkins and Taagepera, 1975). Of course, differences do overemphasize random fluctuations.

Given those difficulties, I worked out the general solution to Richardson's arms race equations. It involves two exponentials (as Richardson noted, without calculating the constants):

$x=Ae^{mt}+Be^{nt}+C$
$y=Ee^{mt}+Fe^{nt}+G$,

where the constants *A* to *G* depend on the original constants (*k, l, a, b, g, h*) and initial conditions in a pretty messy way, as given in Wagner, Perkins and Taagepera (1975). Surprisingly, testing with expanded 1907-1914 data yielded a fit with one single exponential (plus a constant), meaning that the budget increases could be explained as pure self-stimulation, pure mutual stimulation, or varied combinations of the two. In later testing of the "pre-Cuban" Soviet-American race 1945-1960 and the Israeli-Arab race 1949-1972 (Taagepera, Shiffler, Perkins and Wagner, 1975) only the US needed a second exponential -- see Figure 12. This was unsettling.

FIGURE 12. Soviet and United States Military Expenditures, 1945-1960



Source: From Taagepera, Shiffler, Perkins and Wagner (1975). (Curves Show the Richardson Model Fit).

The roof fell in around 1980 when my calculations showed that even the complete solution in two exponentials still included the basic flaw: Bureaucratic self-stimulation could lead to exactly the same observed pattern as mutual stimulation. In other words, even with the best possible data, one could not prove there was a race going on, rather than the military machines operating in isolation and using their existing strength to pressure for ever more resources. This result was never published, while I looked for a conceptual breakthrough, so as to distinguish between mutual stimulation and self-stimulation. None has come.[18]

4.6. Growth of Western Civilization and the Fading of History

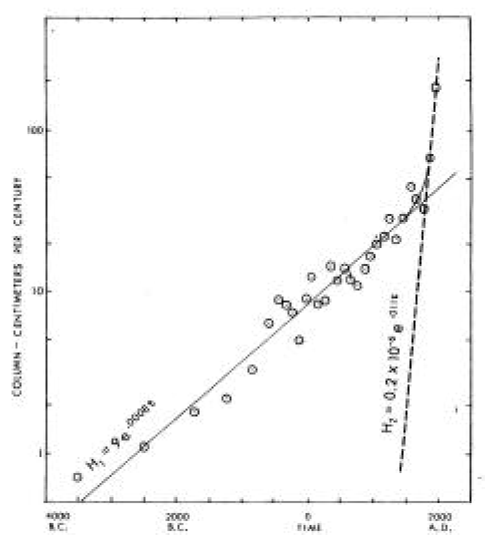
Gray (1966) estimated quantitatively the creativity at various periods of Western history. Recalculating his "creative points" per year (*C*), Taagepera and Colby (1979) found a fair exponential fit from -950 to -400. It is equivalent to $C=3.0x10^{-6}exp(1.0t)$, where t is in **centuries**. This would correspond to the differential equation $dC/dt=1.0C$, meaning an increase rate of 100% per century. More recently (from +885 to +1935), the fit is $C=0.017exp(0.6t)$, meaning $dC/dt=0.6C$ or 60% increase per century. The intermediary period (-400 to +850) offers a decrease, with no clear pattern.

Given the lack of a logical model and the inevitably subjective ratings used as raw data, one can only observe that the exponential pattern occurs over long periods, while failing to occur during an even longer intermediary period. The rate constants during the two exponential periods are similar (100% and 60% per century). We may wonder whether this has something to do with the rate at which existing creativity can foster further increase in creativity -- and possibly a decrease when random chance reduces creativity below a critical mass.

I now note that these rate constants are similar to those of the observed maximal short-term (5-century) decreases of effective number of polities: 90% per century from -950 to -450, and 70% per century from +1500 to +2000. Conversely, the best long-term fit to Gray's creativity points, from -950 to +1935, is around C=0.7exp(0.20t), with a rate constant of 20% per century close to that of N_A, which is 19% per century. There may be a connection.

In part, the apparent growth in creativity may be due to the fading of past history in our minds. Indeed, in nearly all history books and tabulations of important dates more space is given to more recent periods. In the absence of any other information it is reasonable to expect a uniform discount rate as one goes further back in time. This would lead to an exponential pattern.

Figure 13. The Discount Rate of History in CBS News Almanac.



Source: From Taagepera and Colby (1979)

A quick check (Taagepera and Colby 1979) yielded a steep exponential fall for the most recent 500 years (rate constant 1.1 or 110% per century) and a slower one for the preceding 5000 years (0.08 or 8% per century) -- see Figure 13. This is akin to the decay pattern of a mix of two radioactive isotopes. My later analysis of a number of world history books published over two centuries, however, yielded such divergent patterns that this study has remained a conference paper (Taagepera 1988). A main irregularity is that the 19th-century authors often concentrate on Antiquity and give short shrift to modern times.

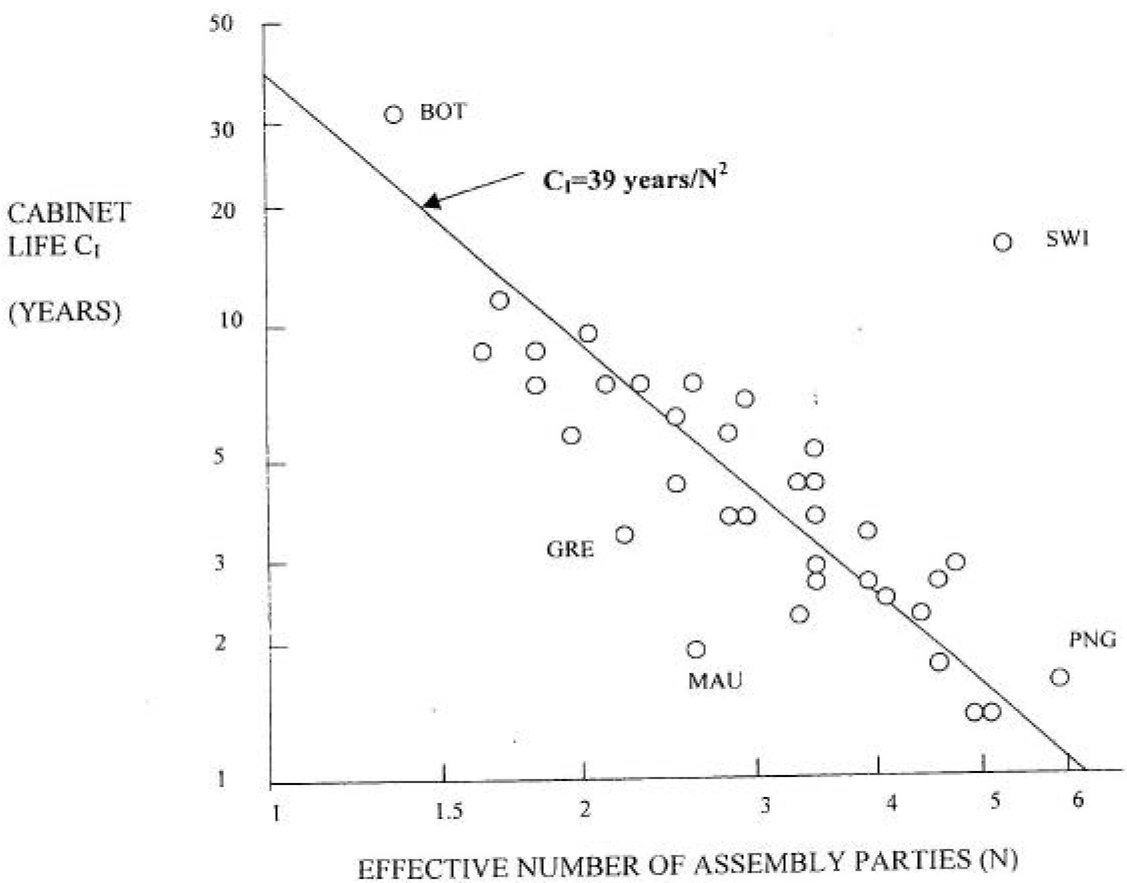
4.7. Inverse Square Law of Coalition Durability

In contrast to previous topics in this section, this one does not deal with growth phenomena expressed as differential equations in time. Instead, it ties the duration (*C*) of a coalition cabinet to the effective number (*N*) of legislative parties: $C=400\ months/N^2$ (Taagepera and Shugart 1989:99-102).[19] A recent recheck with data in Lijphart (1999) yields a higher constant:

$$\mathbf{C = 470\ months/N^2 = 39\ years/N^2}$$

-- see Figure 14. Switzerland with its peculiar cabinet structure is a glaring outlier. Without it, R^2 is around 0.8.

FIGURE 14. Average Cabinet Life (C1) in Years vs. Effective Number of Assemble Parties.



Source: From Taagepera (2002b).
Note: Data from Lijphart (1999: 76-77 and 132-133). BOT=Botswana; GRE=Greece; MAU=Mauritius, PNG=Papua-New Guinea; SWI=Switzerland.

The logical model starts out with the aforementioned observation (Section 3.1 on the cube root law of assembly sizes) that the number of communication channels (c) between p actors is $c=p(p-1)/2$. Even at moderate values of p, $c=p^2/2$ is a fair approximation. If the effective number of parties in an assembly is N, we may assume about N^2 meaningful communication channels among the parties -- and these are also the potential conflict channels. Duration can be expected to be inversely related to the number of conflicts. An inverse square law results. The constant is found empirically, although Taagepera and Shugart (1989) point out why it should be expected to be around 400 months.

Remarkably, by far the best fit is obtained by considering the total number of parties in the assembly as a whole, not just those inside the coalition cabinet. The important implication is that coalition cabinets collapse not only through internal frictions but also through the pull by parties outside the cabinet.

What is the model good for? It cannot be used to predict the duration of a single cabinet, because too many other considerations enter, including random chance. The law applies to averages over long periods. Impressionistically, short cabinet durations in a number of countries have been blamed on an excess of parties, and the empirical tendency has been confirmed by more systematic studies (lately by Lijphart 1999). The inverse square law explains why cabinet duration decreases with increasing number of parties precisely at the rate it does. When institutional changes are considered (such as New Zealand's change in electoral rules in 1996) that can be expected to alter the number of parties represented, the inverse square law allows one to estimate the *average* change in cabinet stability introduced.[20]

4.8. Differential Equations of Cycle of Regimes

This is a model I would lie to construct. It deals with changes over time, yet not in the limiting frames of political games but in a fundamental aspect of the game itself. Aristotle (*Politics*, Book III-15) intimated a cycle of Monarchy --> Oligarchy --> Tyranny --> Democracy --> Monarchy. I would consider a cycle in terms of citizens' and rulers' respective balance of rights and obligations. The differential equations in these four variables should be manageable, given previous examples of prey-predator demographic cycles in biology. The difficulty would be in finding empirical data sufficiently clear-cut for testing and establishing the values of parameters.

5. The Number and Size Distribution of Parties in Assemblies

This section deals with putting specific numbers into Duverger's "law" and "hypothesis" (Duverger 1951, 1954). These are among the most celebrated (and disputed) statements in political science. In the wording used in Taagepera and Shugart (1989: 65): 1) Plurality rule tends to reduce the number of parties to two, regardless of the number of issue dimensions, and 2) proportional representation (PR) rules tend not to reduce the number of parties, if the number of issue dimensions favors the existence of many parties. (For issue dimensions, see Section 5.5.) Since plurality is mostly applied in single-member districts, and several issue dimensions usually exist, the Duverger statements largely boil down to "M=1 leads to two parties, while M>1 leads to three or more parties". Here *M* stands for district magnitude, meaning the number of seats allocated in the district.

Powerful as they are, Duverger's statements represent a semi-quantitative rule rather than a quantitative one. In the case of single-member plurality rule (SMP), how large can third parties be expected to be? How large is the second-largest major party expected to be, compared to the first? In the case of multi-seat PR, how many parties beyond three can one expect, and of what size are the top three parties? Duverger statements offer no answer to such questions. In contrast, this section proceeds to offer specific formulas, predicting specific numbers that can be compared to the actual. In order to reach this goal we have to proceed by small but systematic steps.

5.1. The Number of Seat-winning Parties in District and Nationwide

My direct concern is with parties within an assembly, but much of the argument applies to any set of components that come in integer numbers and add up to a well-defined total.

Consider an electoral district of magnitude *M*, meaning that M seats are to be allocated. How many parties should be expected to win at least one seat? In the absence of any other knowledge, the best guess is that the number of seat-winning parties (*p'*) is square root of *M*:

$$\mathbf{p'=M^{0.5}}$$

(Taagepera and Shugart 1993). If M=1, it's p'=1 (as it should). If M=100, as in The Netherlands prior to 1956, p'=10 -- and this was the case in The Netherlands.[21]

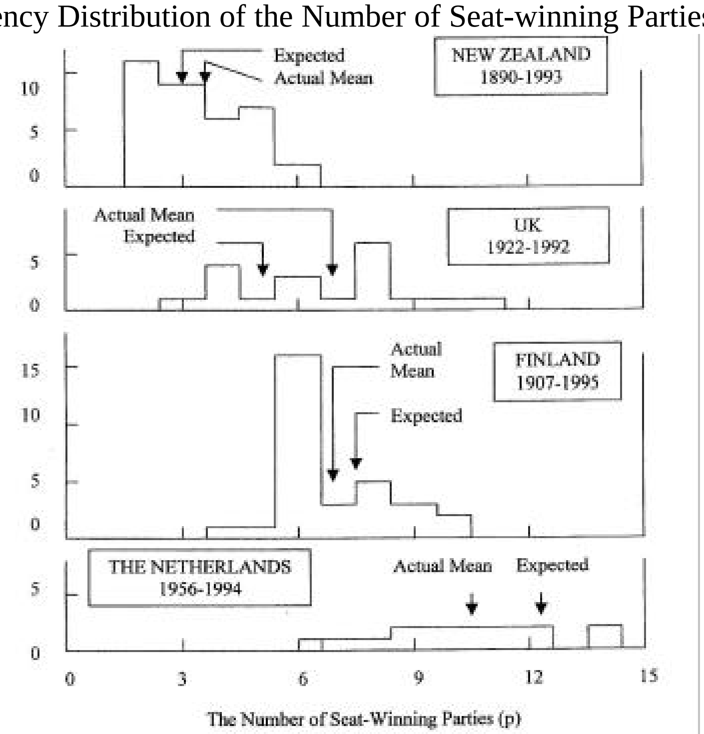
This equation results from an ignorance-based model. The number of seat-winning parties has firm conceptual limits: at least one and at most *M*. The square root of *M* represents the geometric mean of 1 and *M*. Why not use the arithmetic mean, many political scientists keep asking. The quick and dirty answer is: in a district of 100 seats, would you rather expect 10 parties to win seats, or 50? The theoretical reasons are given in Taagepera (1999b).[22]

Nationwide, the number of seat-winning parties (including independents) can be directly estimated only if all seats are distributed within districts of fairly equal magnitude. In this case this number is

$$\mathbf{p=(MS)^{0.25}},$$

where *S* is the total number of seats in the assembly (Taagepera and Shugart 1993). The calculated and actual results for 30 electoral systems are tabulated in Taagepera (2002a), and there is fair agreement.[23] This is the best guess for the average of many elections. For individual elections the number can vary widely, as illustrated in Figure 15.

FIGURE 15. Frequency Distribution of the Number of Seat-winning Parties



Source: From Taagepera (2001b); Expected Value p = (MS)0.25

Once more, the model expressed in the formula above results from considering conceptual limits. If exactly the same parties win seats in all districts, their nationwide number is the same as in a single district: the square root of *M*. This is the lower limit. The upper limit is the number of parties that would win seats, if the entire country formed a single district of magnitude *S*. This leads to square root of *S*. Taking the geometric mean of $M^{0.5}$ and $S^{0.5}$ yields the result above.

What is the use of the result? By itself, a number of parties that includes even the tiniest ones that still manage to win a seat may be of minor interest. But it is the first building block for more to come.

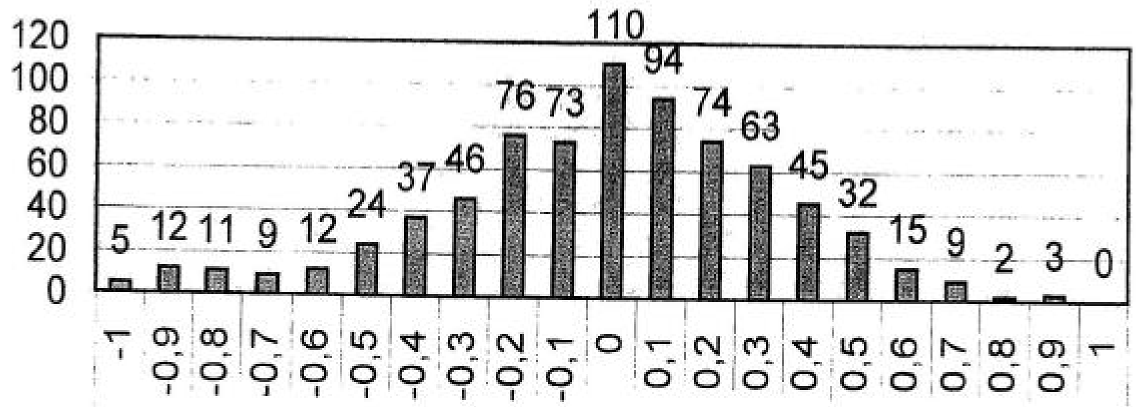
5.2. The Size of the Largest Component

If a well-defined total size *(S)* is divided among *p* parts, our best guess for the size of the largest share (*L*) is the inverse of the square root of the number of parts:

$$\mathbf{L=1/p^{0.5}} \qquad \text{or} \qquad \mathbf{L^2p=1}$$

(Taagepera 1999b). This is a formula of extreme generality. Within 25%, it fits the largest components of federal units in USA, Canada and Australia, both by area and population (Taagepera 1999b). But it can also be tested for every single seat distribution resulting from an election, representing a huge amount of data. The mean of 952 national elections listed in Mackie and Rose (1991, 1997) fits the equation within 1% -- see Figure 16 (Taagepera and Roopalu 2002).[24]

FIGURE 16. Frequency Profile of Normalized L^2p for 752 Elections.



Source: From Taagepera and Roopalu (2002).
Note: *L* is the seat share of the largest party and *p* the number of seatwinning parties. The conceptual bounds on $\log L^2p / \log p$ are -1 (when all shares are equal) and +1 (when the largest share approaches 1). Expected mean and median: 0; actual arithmetic mean: -0.006; actual median: +0.015. Standard deviation: 0.3.

The model again uses the geometric mean of conceptual limits. The largest component cannot be smaller than the mean (*1/p*) nor larger than the whole (1). That's all. While the number of seatwinning parties (previous subsection) depends on district magnitude, the present result does not. Whatever causes the number of components to be what it is, this number has implications for the largest component.

The direct use of the model is in locating anomalies, meaning major deviations from the average pattern. In such cases one might wish to look for the reasons, just as we might do when a distribution deviates from the normal. But this model is also a building block for estimating the entire size structure of parties in an assembly, as seen next.

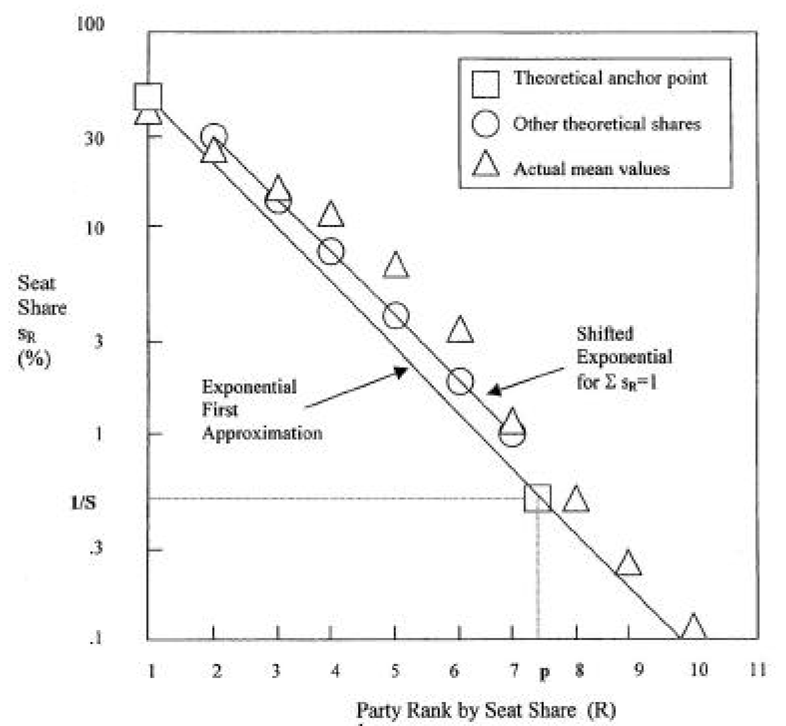
5.3. Distribution of Party Seat Shares

Having the number of seatwinning parties (*p*) and the largest share (*L*), the entire size structure can be estimated, if we can also anchor the other end of the distribution by pinning down the share of the smallest seatwinning party. In the absence of any other knowledge (such as existence of a threshold of votes or seats), this smallest conceivable non-zero share is one seat, or a fraction *1/S* of the total number of seats. Indeed, for electoral systems with no thresholds the observed median number of seats won by the smallest party (or independent) is one seat.

Considerations of continuity would make us place the remaining parties by uniformly decreasing shares between the largest and the smallest. What does uniform decrease from first- to last-ranking party mean? One may consider an arithmetic progression (equal differences, leading to a linear pattern) or geometric progression (equal ratios, leading to an exponential pattern). The arithmetic mean usually leads to the sum of the components being larger than 100 %. The geometric mean most often yields a sum of components lower than 100 %. Thus something in-between may be needed.

Here the model cannot be expressed in a simple equation. Figure 17 shows the shares of parties in a sample country (Finland) graphed versus their rank, on semilog paper (i.e., log of share vs. rank). One way (used in Taagepera 2001b) to estimate the shares of mid-ranked parties is shown there. Place the previously established "anchor points" (first- and last-ranking parties), and draw a straight line between them. Check the resulting total, and shift the line up or down for the intermediary parties, so as to make the sum add up to 100 %.[25]

FIGURE 17. Average Seat Share in Finland, 1907-95, S=200, M=15



Source: From Taagepera (2001b).

Figure 17 shows the typical degree of resemblance between the seat shares thus estimated and the actual long-term averages. On the basis of these estimated seat shares the effective number of assembly parties can be calculated.

5.4. Putting Numbers into Duverger's Law and Hypothesis

Combining the models in the previous subsections, specific numbers can be offered for the sizes of all parties, at least in the case of very simple electoral rules[26] and as averages over many elections. In this sense we are now putting numbers into Duverger's law and hypothesis. The overall flow chart is as follows.

```
S, M ----> p ----> largest share \                    \
                                 ------> other shares  -----> effective number
S ---------------> smallest share /                   /
```

Magnitude and assembly size determine the number of parties. The number of parties determines the largest party's seat share. Assembly size alone determines the smallest party's relative seat share. These two anchor points determine the seat shares of the intermediary parties. These together determine the effective number of parties in the assembly.

Given that each step adds error, a large cumulation of error is to be expected. In the case of four relatively simple systems at extreme values of *M* and *S*, the results are as follows. The combined model yields a good fit for New Zealand, and a fair one for Finland and The Netherlands. For UK the fit is poor but nonetheless reproduces the semiquantitative predictions of Duverger's statements (only two large parties), as shown in the following table (condensed from Taagepera 2001b). This table includes assembly size (*S*), district magnitude (*M*), the number of seatwinning parties (*p*), the seat share of the largest party (*L*) and the effective number of assembly parties (*N*). Subscript *m* indicates prediction by the model, and subscript *a* refers to the actual long-term average value. As one proceeds from low to high *MS*, the actual *p, L* and *N* change in the expected direction.

	S	M	p_m	p_a	L_m	L_a	N_m	N_a
New Zealand 1890-1993	82	1	3.0	3.5	58%	58%	2.0	2.1
UK 1922-1992	629	1	5.0	6.9	45%	56%	2.6	2.2
Finland 1907-1995	200	15	7.4	6.9	37%	34%	3.7	4.5
Netherlands 1956-1994	150	150	12.2	10.6	28%	31%	5.9	4.7

What is the model good for? The brief answer is that whatever the famous Duverger statements (law and hypothesis) are good for, this model is even better for, because it offers more specific figures. The model expresses the pure average effect of district magnitude and assembly size on the seat distribution.

5.5. Issue Dimensions and the Number of Parties

In the wording of Duverger statements at the start of Section 5 the number of issue dimensions (*I*) was mentioned. This refers to the different types of issues on which parties could disagree: socioeconomic, religious, cultural-ethnic, urban-rural, etc. The effective number of parties (*N*) could depend on the issue dimensions according to either of the two following models (Taagepera and Shugart 1989:94):

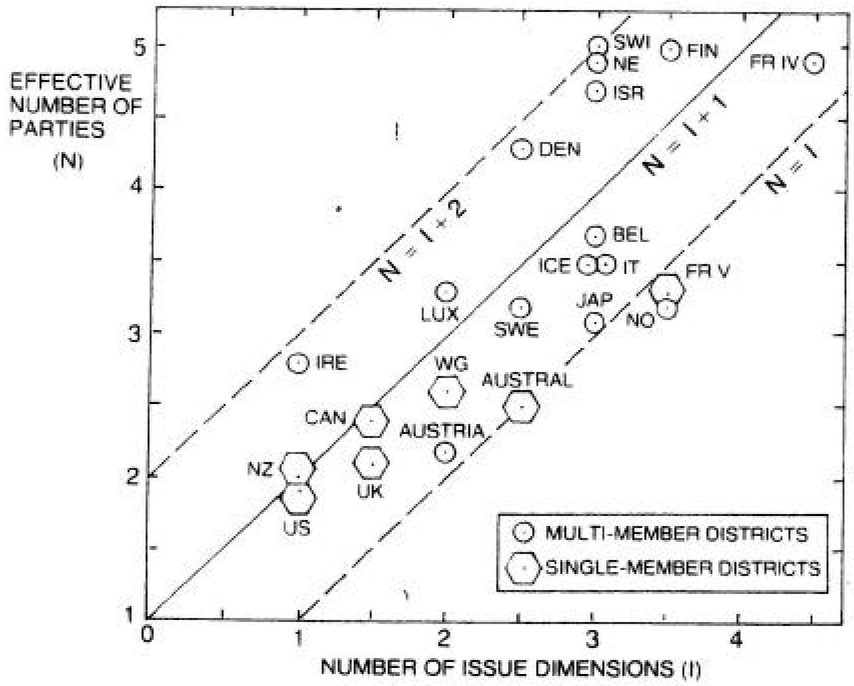
1) Every new issue splits all existing parties. This implies $N=2^I$.
2) Every new issue adds a new party to the existing. This implies $N=I+1$.

The empirical finding (Taagepera and Grofman 1985) is that only one of these models fits the data in Lijphart (1984):

$$\mathbf{N=I+1.}$$

All countries are in the zone ranging from *N=I* to *N=I+2*, and R^2=0.75 -- see Figure 18. Here *N* refers to assembly parties; the fit with electoral parties is poorer, and it is not clear why. More extensive data in Lijphart (1999) yields a steeper best fit than *N=I+1*, but out of 36 countries only Botswana is lower than *I=0*, and only Papua-New-Guinea and Switzerland are above *N=I+2*. So it is too early to scrap the model.

FIGURE 18. The Relationship between N and I in 22 Democracies



Source: From Taagepera and Grofman (1985)

A major problem is that, in contrast to all previous quantities discussed, the number of issue dimensions is not a measured value but a judgment call which itself is heavily influenced by the number of parties. If a social issue does not become a bone of contention among parties, then it may not be counted until it generates a new party. Thus there is a risk of circular reasoning.

The effective number of assembly parties depends on district magnitude on the one hand and on the number of issue dimensions on the other (cf. the wording of Duverger's statements at the start of Section 5). But *N=I+1* ignores *M*, while an expression in *M* derived in Taagepera and Shugart (1993), $N=0.85(MS)^{3/16}$, ignores I. (So do the values of *N* obtained in Section 5.4.) Combining the two, I obtained a somewhat better fit than with each one separately (Taagepera 1999d). This combination,

$$\mathbf{N=I^{0.6}M^{0.15}+1,}$$

satisfies certain boundary conditions but looks complex and artificial. Unless we make progress in operationalizing the measurement of *I*, this equation may not find much use.

6. Converting from People to Seats

This last substantive section involves two separate issues:

1) how are representative assembly seats allocated on the basis of votes; and
2) how are seats in a second chamber or supranational assembly allocated to federal or sovereign states with unequal populations.

Not surprisingly, the two issues are interrelated. Finally, the effect of district magnitude on the vote shares of parties will be modeled, and a graphical way to represent votes-to-seats conversion will be presented and applied.

6.1. The Cube Law and the Seat-vote Equation for M=1

A hundred years ago a regularity in British elections was noted and dubbed the *cube law* -- although it was empirical at the time and hence not a law in the sense of being supported by a logical quantitative model. When v_A and v_B are the vote shares of two parties, and s_A and s_B are the corresponding seat shares allocated by single-member plurality rule (SMP), then the seats ratio tends to be the cube of the votes ratio: $s_A/s_B= (v_A/v_B)^3$. The cube law applies, imperfectly and disputably, to a number of SMP elections for assemblies. For computation of the seat share of one specific party (out of two or more), an equivalent alternative form is more convenient: $s_A=v_A^3/Sv_K^3$, where the summation is over all parties.

More broadly, we might expect that the seats ratio depends on the votes ratio, i.e., $s_A/s_B= f(v_A/v_B)$. It can then be shown (Theil 1969) that the relationship must be of the form

$$\mathbf{s_A/s_B= (v_A/v_B)^n}, \qquad \text{and hence} \qquad \mathbf{s_A=v_A^n /\Sigma v_K^n},$$

where *n* can take any value. This is the only functional form that does not lead to inconsistencies in the presence of more than two parties. But why should *n* be 3?

A logical quantitative model (Taagepera 1973a) specifies that

$$\mathbf{n=logV/logS},$$

where *V* is the total number of votes and *S* the total number of seats. For only two stages (votes and seats), *n* could be a function of total seats and votes in many ways: $n=f(S,V)$. But consistency for three stages (e.g., votes, electoral college, seats) imposes the more specific form $n=f(V)/f(S)$,where $f(V)$ and $f(S)$ have the same functional form. There are reasons to specify further that $f(V)=log\ V$ (see Taagepera 1973a).

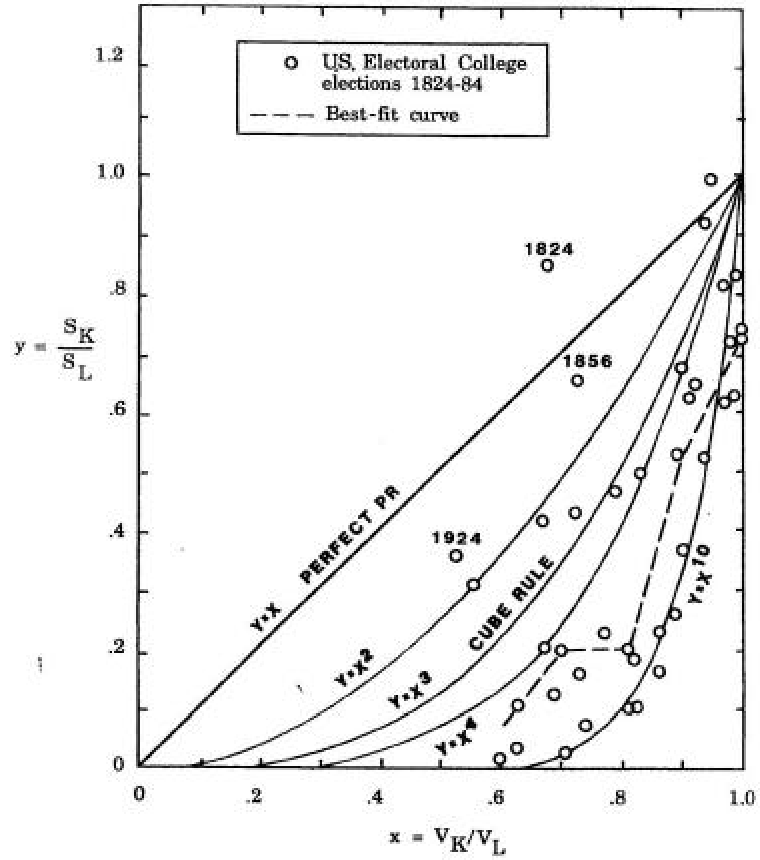
Combining the two equations above leads to what I have called the *seat-vote equation* (SVE):

$$\mathbf{logS\ log(s_A/s_B)= logV\ log(v_A/v_B).}$$

This form is elegantly symmetrical in seats and votes (and possible intermediary stages), but for most purposes it is more practical to use $n=logV/logS$, followed by $s_A=v_A^n /Sv_K^n$. Because of the aforementioned cube root law of assembly sizes ($F=P^{1/3}$, see Section 3.1), *n* is close to 3 in the case of first chambers of representative assemblies. This explains the cube law (and turns it into a true law, based on a combination of two logical models).

But SVE also applies to multi-seat plurality such as used for the US Electoral College. There the best fit to very scattered data is around n=5 -- see Figure 19. In the opposite direction, SVE applies to trade union elections where *n* is as low as 1.5. As extreme boundary cases, SVE also applies to presidential elections (where S=1) and to imaginary elections with as many seats as voters (S=V). In Caribbean countries where the sizes of SMP-elected assemblies fall below the cube root of population, Lijphart (1987) has noted a much steeper seat disparity than expressed by the cube law, in line with SVE.

FIGURE 19. A Case with Minority Attrition Stronger than Predicted by the Cube Rule: U.S. Electoral College Elections



Source: From Taagepera and Shugart (1989: 162).

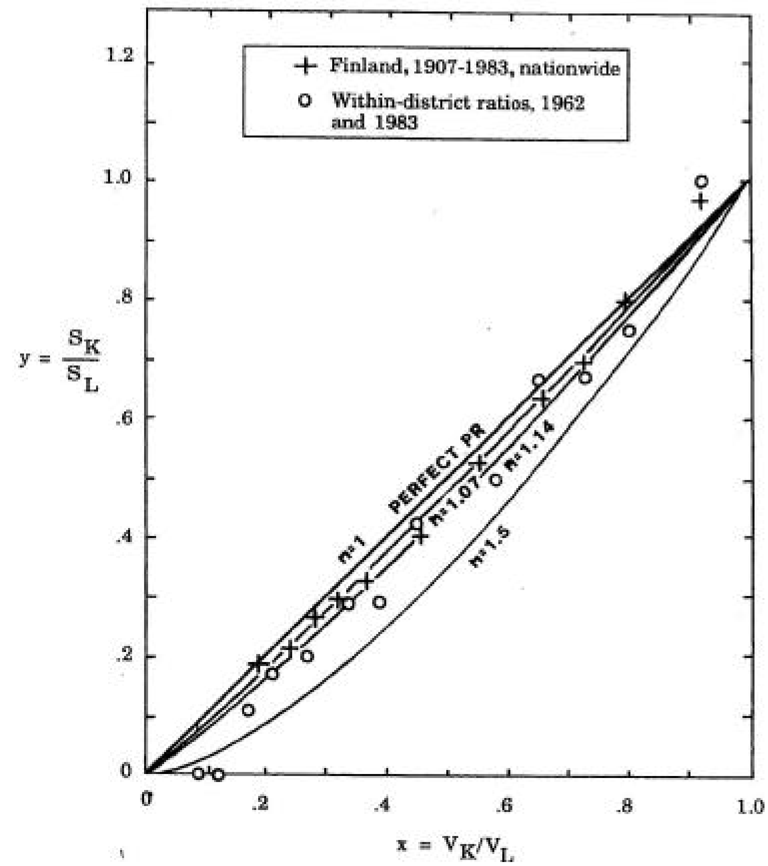
The model building starts by widening the scope beyond SMP. One asks whether the cube law would apply under real or imaginary extreme conditions. Does it apply to presidential elections, where the number of seats is reduced to the conceivable minimum of one seat? The answer is "no" for exponent 3, but "yes" if the exponent were changed to infinity. The opposite extreme case is the imaginary situation where the elected assembly is as large as possible, namely including all voters.[27] This would lead to perfect proportionality, requiring an exponent "one". These considerations expand $s_A/s_B= (v_A/v_B)^3$ to $s_A/s_B= (v_A/v_B)^n$, in line with Theil (1969).

Compared to the cube law, this may look at the first glance nihilistic: instead of n=3, any value of exponent *n* goes. However, assuming continuity, the extreme cases also say that as *S* expands from 1 to *V*, the exponent *n* must decrease from infinity to one. It so happens that $n=logV/logS$ is the simplest way to satisfy these boundary and continuity conditions -- and it also satisfies certain other desirable criteria. [28]

6.2. The Seat-vote Equation for Multiseat PR Elections

Expansion of SVE to multiseat PR elections owes to George Shiffler, the undergraduate who also co-authored an arms race article (Taagepera, Shiffler, Perkins and Wagner, 1975). Around 1975 he boldly proposed that with PR in districts of magnitude *M*, the ratio of logarithms in $n=logV/logS$ should be put to the power $1/M$. With logV/logS=3 (cube root law) and M=15 (as in Finland), $n=(logV/logS)^{1/M}$ leads to n=1.07 -- hardly distinct from the conceptual limit of n=1 (Taagepera and Shugart 1989:186) -- see Figure 20.

FIGURE 20. Seat Ratios and Vote Ratios in Near-PR Systems: Finland and Individual Finnish Districts.



Source: From Taagepera and Shugart (1989: 168).
Note: The graph shows medians of groups of 11 data points closest to each other on the x-scale.

For years, I discarded this idea. First, there was no logical model behind the expression. Second, I mistakenly thought the true expression should go to n=1 when M=S. Third, there was no way to test such small deviations from n=1. But when I developed more refined ways to test the seat-vote relationships (Taagepera 1986), George turned out to be in the right ball park. Reed (1996) has carried out a detailed testing of SVE regarding Japan, where district magnitudes are low and hence the resulting n=1.3 is relatively high as far as PR goes.

The only remaining specification needed (Taagepera 1986) was to include multi-seat plurality elections by replacing district magnitude *M* by "allocation magnitude" *M'*, which is the number of parties that conceivably could obtain seats in the given district. For PR, *M'* is simply *M*, but for multiseat plurality M'=1 for any *M*. So the general expression in PR *and* plurality elections is

$$\mathbf{n=(logV/logS)^{1/M'}}.$$

It still nags me that the power *1/M* does not have a logical foundation, in contrast to SVE for plurality elections. Yes, it does satisfy the constraint that for M'=1 the general equation is reduced to the specific one (for plurality). And it also seems to reflect the reality for PR (although testing is difficult and has been limited). But, simplicity apart, why should the exponent be *1/M* rather than, say, $(1/M)^2$ or $(1/M)^{0.5}$? Maybe another curious undergraduate will supply the answer.

6.3. Seat Allocation in Second Chambers and EU Assemblies: The Seat-population Equation

This issue is highly topical because of the expansion of the European Union (EU). How many seats should various countries receive in the European Parliament (EP)? And how many "votes" should they receive in the Council of the EU (CEU)? What makes the answer difficult is that the countries must be represented both as being equal (since they are distinct entities) and also as unequal (given huge disparities in populations). The same problem arises for some upper chamber (such as Canada's). One extreme is to represent all entities equally and ignore disparities in population (e.g., the US Senate and the UN General Assembly). The other extreme approach would be to represent populations proportionately; but then Luxembourg would obtain no seats at all in the EU bodies.

The compromise is to over-represent the smaller countries while still giving them fewer seats than to larger countries. In terms of SVE format, it means a value of exponent *n* smaller than 1. Indeed, the following "*seat-population equation*" emerges from boundary conditions (Taagepera and Hosli 2002). For two countries (or federal subunits) A and B, this equation connects their seat shares (or "vote" shares in the CEU), s_A and s_B, to their populations (p_A and p_B), taking into account the total number of countries (N), their total population (P) and the total number of seats (S):

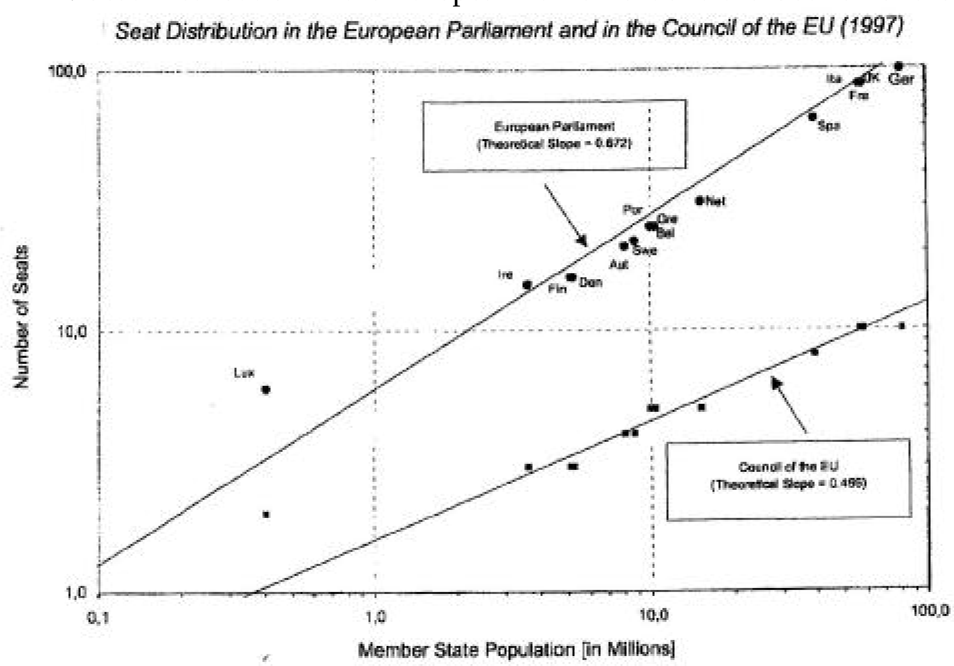
$$\mathbf{s_A/s_B = (p_A/p_B)^n} \qquad \text{or} \qquad \mathbf{s_A = p_A^n / \Sigma p_K^n},$$

where

$$\mathbf{n=(1/logN-1/logS)/(1/logN-1/logP).}$$

In 1997, EU had 365 million people in 15 countries. EP (626 seats) yields n=0.672, while CEU (87 "votes") yields n=0.456, meaning a steeper overrepresentation of smaller countries. The agreement with the successive EP seat and CEU "vote" distributions from 1970 to 1997 is remarkable (except for heavy actual overrepresentation of Luxembourg, compared to the model) -- see Figure 21.

FIGURE 21. Seat Distribution in the European Parliament and the Council of the EU (1997)



Source: From Taagepera and Hosli (2002).

Various empirical fits for CEU seat distributions have been proposed previously. (Surprisingly, none have for EP.) They include the use of SVE with the exponent *n* arbitrarily set at 0.5. The beauty of the new "seat-population equation" is that the *same* equation applies to both EU bodies, without the need to insert any empirical coefficients. It is not a retroactive data fit, a postdiction. It is truly predictive, once *N*, *S* and the country populations are given.

The logical quantitative model uses three constraints or boundary conditions. The first two parallel the SVE:

1) For only one seat available, the exponent *n* in $s_A/s_B=(p_A/p_B)^n$ must tend to infinity, allocating the single seat to the largest country.
2) For as many seats as the total population, we must have n=1.
3) The new additional constraint comes in when there are as many seats as there are countries; then one seat should be assigned to each country, and n=0 yields this outcome.

The seat-population equation satisfies these constraints. It may not be unique doing so, but it is the simplest way.

Implications for the proposed expansion of EU come on two levels. First, the optimal total sizes of EP and CEU are suggested by the aforementioned extension of the cube root law (Taagepera and Recchia 2002 -- see Section 3.2). Second, either this optimal size or the actual size chosen by EU can be used as a starting point for allocating the seats among the countries through the seat-population equation. The actual allocation has fallen in line with this equation for the last 30 years. The changes proposed in Nice, December 2000, however, tend to deviate from the equation (Taagepera and Hosli 2002), which means they may be excessively influenced by political games. Application of the model to second chambers such as Canada's remains to be completed.

6.4. Distribution of Party Vote Shares

In Section 5.4 the distribution of party seat shares was modeled, thus putting specific numbers into Duverger's statements. By reversing the SVE we can use these expected seat shares to calculate the corresponding vote shares:

$$\mathbf{v_A/v_B=(s_A/s_B)^{1/n} \quad \text{or} \quad v_A=s_A^{\,n} / \Sigma s_K^{\,1/n},}$$

where

$$\mathbf{n=(\log V/\log S)^{1/M'}.}$$

On this basis the effective number of electoral parties can be calculated.

This completes putting numbers into Duverger's statements. Compared to the effect of district magnitude and assembly size on seat shares, their effect on vote shares is more indirect. The model involves further simplifying assumptions; therefore, further deviations can be expected. Still, there is some resemblance between the model-based and actual vote shares of the largest party and of the effective number of parties.[29] Using the same countries as in Section 5.4 (based on Taagepera 2001b), the picture is as follows. In the following table, *L* is the vote share of the largest party, and *N* is the effective number of parties based on votes. Subscript *m* indicates prediction by the model, and subscript *a* refers to the actual average value. Only UK is out of line in a major way.

	S	M	L_m	L_a	N_m	N_a
New Zealand 1890-1993	82	1	42%	47%	2.9	2.7
UK 1922-1992	629	1	*32%*	45%	*4.0*	2.8
Finland 1907-1995	200	15	35%	32%	4.0	4.9
Netherlands 1956-1994	150	150	28%	30%	6.1	4.9

6.5. Prediction of Proportionality Profiles

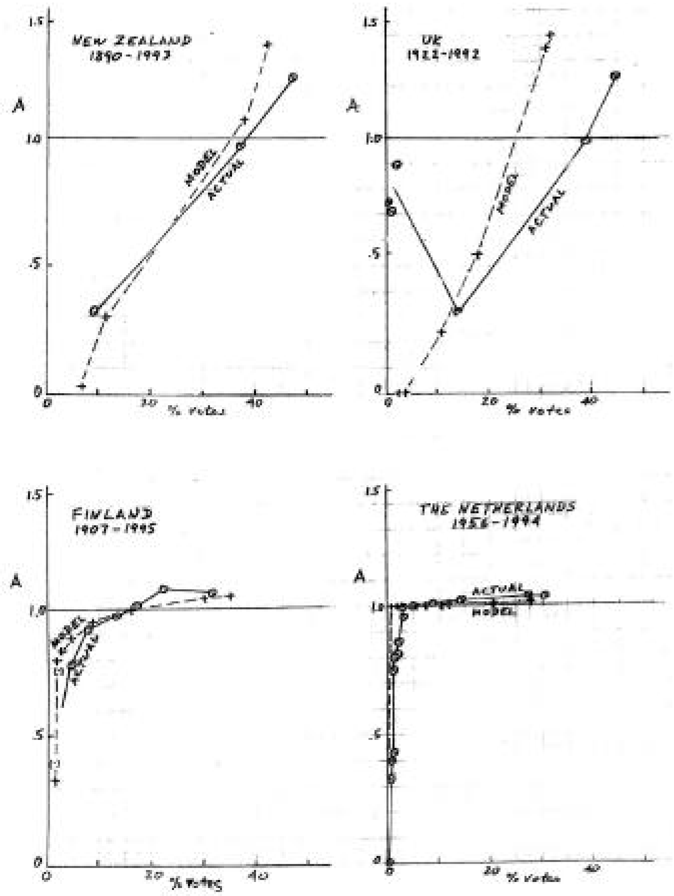
It remains to visualize the average long-term seat-vote patterns predicted by the models and compare them with the actual patterns. The following format brings best out the fine differences of various empirical and theoretical patterns. For each party, calculate the "*advantage ratio*"

A=%seats/%votes,

and graph it against the % votes (Taagepera and Laakso 1980; Taagepera and Shugart 1989:67-76). Because perfect proportionality corresponds to the horizontal line at A=1, deviations from proportionality are brought out most visibly. This approach can be applied to purely empirical data or for comparing models with data.

Earlier testing of the seat-vote equation (Taagepera 1986; Taagepera and Shugart 1989:191-196) still fell short of a prediction on purely institutional grounds because the average effective number of electoral parties was among the inputs. Taagepera (2001b) goes essentially all the way: For seat shares only the institutional inputs of assembly size (*S*) and district magnitude (*M*) are used. For vote shares (and hence the advantage ratio) the total number (*V*) of voters also enters in *n=logV/logS*. To the extent the cube root law of assembly sizes determines *S,* the total input is reduced to *M* and *V*.

FIGURE 22. Proportionality Profiles of Countries with Different Combinations of Assembly Size and District Magnitude: Model and Actual Mean Values for Largest Party, Second-largest, etc.



Source: Based on data in Taagepera (2001b). A=% seats/% votes.

Figure 22 shows theoretical and actual proportionality profiles for four countries with widely different combinations of *M* and *S* (based on Taagepera 2001b). The agreement of the model with the empirical averages is fair for The Netherlands, Finland and also New Zealand (except for overestimating the largest party advantage in New Zealand). UK is appreciably off the mark. Many more countries need to be tested, and this should be my focus during the next few years.

7. Review of Types of Models

What types of model-building methods have been used here? I'll try systematize the various approaches in hindsight. While solving a particular problem, one of course catches it as catch can, sometimes shifting from one approach to another. I'll try to point out such mixed strategies too.

7.1. Boundary Conditions as Anchor Points

Boundary conditions are a form of constraints -- conceptual extreme values beyond which a quantity cannot go. Sometimes these extremes can occur in practice (such as all parties having the same number of seats, in Section 5.2); at other times they can be approached but not reached (such as the largest party having all the seats, leaving nothing to other seatwinning parties). As a first approximation we can overlook this distinction. The main thing is that values outside the extremes are impossible not only practically but also conceptually. Compare the assertion of having caught a fish *plus* 10 feet long versus one *minus* 10 feet long. Boundary conditions can be used to estimate the median value likely to be encountered. In conjunction with assumption of continuity, they can anchor a set of varying values. The following comments deal with some subvarieties.

The geometric mean of conceptual extremes: The model of almost complete ignorance. If the conceptual extremes are all we know, then the best we can do is to guess that values around the median will materialize, since we have no reason to shift away from the median. When both conceptual extremes are positive and of different orders of magnitude, the median tends to be around the geometric mean of the extremes (Taagepera 1999b).

Sections 3.2 and 3.3 apply this approach to federally constituted second chambers, where the extreme choices are whether to represent only the territorial units or only the total population. Here both extremes actually occur with appreciable frequency, but the median still is around the geometric mean. Section 4.3 sets two boundaries for the effective number of polities by population: one single polity, and as many as there are on area basis. Here the boundary values are not observed to occur. The geometric mean is a fair fit over 5000 years.

Section 5.1 sets the number of seatwinning parties between 1 and *M* in a single district, and between $M^{0.5}$ and $S^{0.5}$ nationwide. In the district the boundary values are frequently reached when the district magnitude is very low. For the nationwide, the boundaries no longer are true conceptual boundaries but previous best guesses at medians. And this cumulating may continue in Section 5.2, where the lower boundary on the largest party's seat share may be based either on the actual number of seatwinning parties or on the theoretical estimate. (The upper boundary always is the total number of seats.) The theoretical estimate in Section 5.2 is based on a reachable lower boundary (all seatwinning parties have equal shares) and on an unreachable upper one (the largest party has all seats).

Continuity between extreme ranks. In contrast to the previous situation, we may have the extra information that the anchor points are populated and that all other cases must fall in-between. This is by definition the case when dealing with rankings. With the population of the first- and last-ranking components known, the best we can do in the absence of any further information is to guess at quasi-continuous decrease in-between. Two different cases arise: either we know the total number of components, or we don't.

1) For seat shares of parties an estimate of the total number of seatwinning parties is known from the geometric mean of conceptual extremes. It's a matter of spreading the sizes of the known number of

intermediary parties uniformly between the largest and the smallest, subject to a fixed total. Sections 5.3 and 5.4 find that the actual patterns range from exponential to linear.

2) For ranked cities, the total number of entities is not known. The size of the largest component is known, and a simple relationship with rank is empirically observed. In Section 3.5 the strategy is to define logically another anchor at the low end (a settlement of one person) and assume that the observed relationship continues to the very end. A figure for the total population emerges and can be compared to the actual. The rankings of journals (Section 3.6) follow the same pattern, and here the low anchor (at one item) is very real and documented, in contrast to its artificial nature in the case of cities.

Continuity between extremes on two quasi-continuous scales. In the seat-vote equation (SVE, Section 6.1, leading to 6.4 and 6.5) the independent variable is different from the previous in that discontinuous rankings are replaced by a quasi-continuous variable: the votes. Rankings impose a non-positive slope, but this is not the case here. The boundary conditions on exponent n are the following: $S=V$ leads to n=1; $S=1$ leads to n-->infinity. A further requirement (to be discussed later) is applicability to more than two stages. Assuming continuity and maximum simplicity, $n=\log V/\log S$ results. Extension to multiseat election (Section 6.2) follows the same approach but is incomplete theoretically because only one anchor is defined: For M=1 the expression for n must reduce itself to a previous form. The rest is based on picking the simplest form possible and finding that it fits the data.

Continuity between extremes on three quasi-continuous scales. For seat allocation in EU assemblies and federal chambers that pay attention to different populations of constituent units (Section 6.3), exponent n must satisfy boundary conditions for two independent variables, the number of countries and their total population. $S=P$ leads to n=1; $S=1$ leads to n-->infinity; and $S=N$ leads to n=0. The simplest mathematical form that satisfies these constraints turns out to fit the data. Except for increased complexity, the approach is the same as for SVE.

7.2. Changes in Time and Space

Assuming continuity or at least quasi-continuity (as in the case of votes or electric current, carried by discrete electrons), differential equations are the standard way to deduce the overall picture from what happens at one instant.

Ignorance about the direction of change. If we have no idea whether a quantity is increasing or decreasing over time (or a space direction), the best possible assumption is that it remains the same. Any other assumption presumes some further information. If we have the further information that something increases on its own (rather than being built from the outside), the best possible assumption is that the relative growth rate of this variable remains constant. The solution of the resulting differential equation is the exponential function. The rate constants are determined empirically, in the absence of other information. In this sense, the exponential model is an ignorance-based one.

A single variable dependent on time. This is the case for the fading of history and the two growth periods of Western civilization (Section 4.5). The exponential model of ignorance fits to a limited extent. The empirically determined rate constants offer possibilities for comparison.

Two dependent variables with non-time-dependent interconnection. The basic assumptions and solutions are the same as above. However, effective numbers of polities in terms of population and area (Section 4.3, leading to 4.4) are interconnected at any instant in a non-time-dependent way, based on a model of ignorance: The rate constant of the first is one-half of that of the second. Otherwise the rate constants are empirical.

A system of two equations in two dependent variables. The change rate of each variable depends on the values of both variables. In arms races (Section 4.6) the impacts of both variables add, while for world population growth (Section 4.1) they multiply. They also multiply, in a way, in hyperinflation (Section 4.2). Solutions are correspondingly more complex.

Changes in space. The model for trade/GNP ratio (Section 3.4) starts with the assumption of constant absorption rate over distance. In one dimension, this leads to a simple differential equation with exponential solution. However, in contrast to time, space has multiple dimensions, adding complexity. As in physics, extension from one to several dimensions involves technical difficulties and approximate solutions. Fortunately, the trading space is only two-dimensional. Complexity is compounded by insertion of constructs called borders.

7.3. The Number of Communication Channels

This may be a key variable in many social issues, given that society means communication. I have used the number of communication channels in two very different ways. In duration of cabinet coalitions (Section 4.7) this number is plugged in directly, as it is. In contrast, the minimum value of this number is sought when determining the size of representative assemblies (Section 3.1). The latter has implications for the sizes of second chambers and European assemblies (Sections 3.2 and 3.3), more remotely for the distribution of seat shares in all these bodies (all of Section 5), and even more remotely on the vote shares of parties (Sections 6.4 and 6.5).

7.4. Minimization and Maximization

Minimization or maximization of some quantity (implying dy/dx=0) is a frequent approach in physics but has remained relatively rare in political science, ever since Kochen and Deutsch (1969) used it for a model of optimal decentralization. Of course, Richardson (1960) used dx/dt=0 for placing the stability lines for arms races. In retrospect, I'm surprised how little I have used it myself, apart from assembly sizes (Section 3.1).

7.5. Consistency with More Than Two Components or Stages

If something explains the relation between A and B and also A and C, does it apply between B and C? It must, if it is to be general. The seat-vote equation (Section 6.1) uses this requirement twice. Once one accepts that the seat ratio s_A/s_B of two parties must be some function $f(v_A/v_B)$ of their vote ratio, a profusion of functional forms present themselves. However, as Henri Theil (1969) shows, when one considers more than two parties, only one form guarantees internal consistency, and that is $s_A/s_B = (v_A/v_B)^n$.

In determining the exponent *n* in the previous equation, analogous considerations enter, but with stages instead of components. For only two stages (votes and seats), *n* could be a function of total seats and votes in many ways: $n=n(S,V)$. But consistency for three stages (e.g., votes, electoral college, seats) imposes that *n* be a function of *V* divided by the same function of *S*: $n=f(V)/f(S)$. Why $f(V)$ should be $logV$ is another matter.

Three-stage elections are rare, but a universal model must be consistent with them -- and this constraint narrows down our choice of functional relationships in a fruitful way. Consistency with more than two components or stages is a question to be posed whenever one constructs a relationship between two quantities: Does it lead to consistency when a third quantity of the same type is added?

7.6. Replacing Unequal Shares with Equal Shares

This is the task when defining the effective number of components (Section 2). We must have $Ss_i=1$ by definition when the summation is over all shares. But $Ss_i^2<1$. If we assign meaning to the squares of shares (as we do, for instance, in least square calculations), we may ask which number (*N*) of equal shares would yield the same sum of squares as do the actual components. Those equal shares still must add up to 1. Hence each share is $1/N$, and its square is $1/N^2$. This means that $N(1/N)^2=Ss_i^2$, so that $N=1/s_i^2$ results as the effective number of components.

Of course, this is the effective number only in the sense of yielding the same sum of squares as the original constellation. If we want, say, the same sum of cubes, the effective number is different. The general formula for Ss_i^n with any n is given in Laakso and Taagepera (1979) and Taagepera and Shugart (1989:259). However, n=2 is the simplest and has some other desirable features. In particular, it represents self-weighting: each share is assigned a weight equal to itself.

Replacing of unequal shares with equal shares -- is this model building or mere definition of a new quantity? It's in-between. The notion of equal shares that are somehow equivalent to the actual shares is akin to a model. However, a model must be testable and may be disproved by empirical data. In contrast, the effective number of components always works out in a formal way. Thus it is a definition grounded in the spirit of model building.[30]

7.7. Addition or Multiplication of Components

This is a broad issue discussed in Taagepera and Shugart (1989:243) and Taagepera (1999b). I'd say that physicists instinctively multiply while social scientists instinctively add. Consider the following example. To remain alive we need air (A) and food (F). The product AF implies that no amount of food can replace air. In contrast, an addition a_1A+a_2F would suggest that plenty of food can substitute for zero air. I remain a physicist: In case of doubt, better multiply than add.

In the case of parties and issue dimensions (Section 5.5), however, addition turns out to be the proper course. The model $N=I+1$ fits data, while the multiplicative $N=2^I$ does not. But note that this is not a knee-jerk addition a_1I+a_2, with contentless coefficients. There are logical reasons why we should have $a_1=a_2=1$.

7.8. Mixed Approaches

Of course, mixing and sequences occur. The effective number of polities (Section 4.3) starts with geometric mean of conceptual extremes to connect area and population, and then continues with the exponential assumption. The number and size distribution of parties (Section 5) applies geometric means of conceptual extremes over and over, and then throws in a quasi-exponential distribution. The seat-vote equation (Section 6.1) results from considerations of conceptual extremes, plus consistency with more than two components and stages. The same is the case for the seat-population equation (Section 6.3).

8. Conclusion

I have presented here some methods to construct models that are logical and quantitative. Various testable predictions result regarding the framework surrounding politics. These methods are not a panacea. But they come handy in addressing certain types of problems among those of interest in the study of politics.

Bibliography

Blondel, Jean (1969), *Introduction to Comparative Government.* New York: Praeger.
Cagan, Phillip (1956), The monetary dynamics of hyperinflation. In Milton Friedman, ed., *Studies in the Quantity Theory of Money,* University of Chicago Press.
Cox, Gary (1997), *Making Votes Count.* Cambridge: Cambridge University Press.
Dahl, Robert A., and Edward R. Tufte (1973), *Size and Democracy*. Stanford: Stanford University Press.
Duverger, Maurice (1951), *Les partis politiques.* Paris: Le Seuil.
Duverger, Maurice (1954), *Political Parties: Their Organization and Activity in the Modern State.* New York: Wiley.
Feld, Scott, and Bernard Grofman (1977), Variation in class size, the class size paradox and some consequences for students, *Research in Higher Education* 6, 215-222.
Finifter, Ada W., ed. (1993), *Political Science: The State of the Discipline II.* Washington, DC: The American Political Science Association.
Goodin, Robert E., and Hans-Dieter Klingemann, eds. (1996), *A New Handbook of Political Science.* Oxford and New York: Oxford University Press.
Gray, Charles E. (1966), A measurement of creativity in Western civilization, *American Anthropologist* 68, 1384-1417.
Hart, Hornell (1945), Logistic social trends, *American Journal of Sociology* 50, 337-352.
Kochen, Manfred, and Karl W. Deutsch (1969), Toward a rational theory of decentralization, *American Political Science Review* 63, 734.
Laakso, Markku, and Rein Taagepera (1979), 'Effective' number of parties: A measure with application to West Europe, *Comparative Political Studies* 23, 3-27.
Lanchester, F.L. (1956), Mathematics in warfare. In J.R. Newman, ed., *The World of Mathematics,* vol. 4, 2136.
Lijphart, Arend (1984), *Democracies: Patterns of Majoritarian and Consensus Government in Twenty-One Countries.* New Haven and London: Yale University Press.
Lijphart, Arend (1987), Size, pluralism, and the Westminster model of democracy: Implications for the Eastern Caribbean. In Jorge Heine, ed., *A Revolution Aborted: The Lessons of Grenada.* Pittsburgh: University of Pittsburgh Press, 321-340.
Lijphart, Arend (1994), *Electoral Systems and Party Systems.* Oxford: Oxford University Press.
Lijphart, Arend (1999), *Patterns of Democracy: Government Forms and Performance in Thirty-Six Countries.* New Haven: Yale University Press.
Mackie, Thomas T. and Richard Rose (1991), *The International Almanac of Electoral History.* London: Macmillan, and Washington: Congressional Quarterly.
Mackie, Thomas T. and Richard Rose (1997), *A Decade of Election Results: Updating the International Almanac.* Glasgow: Centre for the Study of Public Policy, University of Strathclyde.
Meyer, François (1974), *La surchauffe de la croissance.* Paris: Fayard.
Misiunas, Romuald, and Rein Taagepera (1993), *The Baltic States: Years of Dependence 1940-1990.* London: Hurst, and Berkeley: University of California Press. Also published in Lithuanian (1992), Hungarian (1994), and Estonian (1997).
Rae, Douglas W. (1967), *The Political Consequences of Electoral Laws.* New Haven: Yale University Press. Second edition in 1971.
Reed, Steven R. (1996) Seats and votes: Testing Taagepera in Japan, *Electoral Studies* 5, 71-81.
Richardson, Lewis F. (1960), *Arms and Insecurity,* Pittsburgh: Boxwood.
Taagepera, Rein (1968), Growth curves of empires, *General Systems* 13, 171-175.
Taagepera, Rein (1972), The size of national assemblies, *Social Science Research* 1:4, 385-401.
Taagepera, Rein (1973a), Seats and votes: A generalization of the cube law of elections, *Social Science Research* 2:3, 257-275.
Taagepera, Rein (1973b), Fractional iteration of exp x, and fractional arithmetic operations. Social Sciences Working Paper 44c, University of California, Irvine.
Taagepera, Rein (1976a), Why the trade/GNP ratio decreases with country size, *Social Science*

Research 5, 385-404.
Taagepera, Rein (1976b), Crisis around 2005 AD? A technology-population interaction model, *General Systems* 21, 137-138.
Taagepera, Rein (1978a), Size and duration of empires: Systematics of size, *Social Science Research* 7, 108-127.
Taagepera, Rein (1978b), Size and duration of empires: Growth-decline curves, 3000 to 600 BC, *Social Science Research* 7, 180-196.
Taagepera, Rein (1979a), Inequality, concentration, imbalance, *Political Methodology* 6, 275-291.
Taagepera, Rein (1979b), People, skills and resources: An interaction model for world population growth, *Technological Forecasting and Social Change* 13, 13-30.
Taagepera, Rein (1979c), Size and duration of empires: Growth-decline curves, 600 BC to 600 AD, *Social Science History* 3, 115-139.
Taagepera, Rein (1979-80), Stockpile-budget and ratio interaction models for arms races, *Papers* of the Peace Science Society (International) 29, 67-78.
Taagepera, Rein (1981), Population crisis and the Baltics, *Journal of Baltic Studies* 12, 234-244.
Taagepera, Rein (1984), *Softening without Liberalization in the Soviet Union: The Case of Jüri Kukk.* Lanham, MD: University Press of America.
Taagepera, Rein (1986), Reformulating the cube law of elections for proportional representation elections, *American Political Science Review* 80, 489-504.
Taagepera, Rein (1987), The inverse square law of coalition durability. International Conference on Coalition Theory and Public Choice, Fiesole (Italy), 25-29 May.
Taagepera, Rein (1988), The fading rate of history. 17th Annual Meeting of the International Society for the Comparative Study of Civilization, Hampton University, May 26-29.
Taagepera, Rein (1989), Empirical threshold of representation, *Electoral Studies* 8, 105-116.
Taagepera, Rein (1993), *Estonia: Return to Independence.* Boulder, CO: Westview Press.
Taagepera, Rein (1997a), Effective number of parties for incomplete data, *Electoral Studies* 16, 145-151.
Taagepera, Rein (1997b), Expansion and contraction patterns of large polities: Context for Russia, *International Studies Quarterly* 41, 475-504.
Taagepera, Rein (1998a), Effective magnitude and effective threshold, *Electoral Studies* 17:4, 393-404.
Taagepera, Rein (1998b), Nationwide inclusion and exclusion thresholds of representation, *Electoral Studies* 17:4, 405-417.
Taagepera, Rein (1999a), *The Finno-Ugric Republics and the Russian State.* London: Hurst. Also published in Estonian (2000) and Hungarian (2000).
Taagepera, Rein (1999b), Ignorance-based quantitative models and their practical implications, *Journal of Theoretical Politics* 11:3, 421-431.
Taagepera, Rein (1999c), Supplementing the effective number of parties, *Electoral Studies* 18:4, 497-504.
Taagepera, Rein (1999d), The number of parties as a function of heterogeneity and electoral system, *Comparative Political Studies* 32:5, 531-548.
Taagepera, Rein (2000), Should Russia break up. Annual Meeting of the American Political Science Association, Washington, DC, 31 August - 3 September 2000. Also Research Monograph # 36, Center for the Study of Democracy, University of California, Irvine (November 1999) www.democ.uci.edu/democ .
Taagepera, Rein (2001a), Should Russia break up. Conference on Nationality and Citizenship in Post-Communist Europe, Paris, 9-7 July.
Taagepera, Rein (2001b), Party size baselines imposed by institutional constraints: theory for simple electoral systems, *Journal of Theoretical Politics* 13, 331-354.
Taagepera, Rein (2001c), Journal citation frequency: The Hyperbolic pattern. Research Monograph CSD01-08, Center for the Study of Democracy, University of California (November 2001) www.democ.uci.edu/democ .
Taagepera, Rein (2001d), The logical underpinnings of the hyperbolic rank-size rule. Manuscript.
Taagepera, Rein (2002a), Nationwide effective threshold of representation, *Electoral Studies*, forthcoming.

Taagepera, Rein (2002b), Arend Lijphart's dimensions of democracy: Logical connections and institutional design. Manuscript.
Taagepera, Rein, and Benjamin N. Colby (1979), Growth of western civilization: Epicyclical or exponential? *American Anthropologist* 4, 907-912.
Taagepera, Rein, and Bernard Grofman (1981), Effective size and number of components, *Sociological Methods and Research* 10, 63-81.
Taagepera, Rein, and Bernard Grofman (1985), Rethinking Duverger's law: Predicting the effective number of parties in plurality and PR systems -- Parties minus issues equals one, *European Journal of Political Research* 13, 341-352.
Taagepera, Rein, and James P. Hayes (1977), How trade/GNP ratio decreases with country size, *Social Science Research* 6, 108-132.
Taagepera, Rein, and Madeleine O. Hosli (2002), National representation in international assemblies: A seat distribution formula for the EU Council and Parliament. Manuscript.
Taagepera, Rein, and Edgar Kaskla (2001), The city-country rule: An Extension of the rank-size rule, *Journal of World-Systems Research* 7, 157-174.
Taagepera, Rein, and Markku Laakso (1980), Proportionality profiles of West European electoral systems, *European Journal of Political Research* 8, 423-446.
Taagepera, Rein, and Matti Nurmia (1961), On the relations between half-life and energy release in alpha-decay, *Ann. Acad. Sci. Fennicae* A. VI. 78. Referred to in Hyde, Pearlman and Seaborg, *Nuclear properties of the heavy elements, 1*, 256-58, 1964, and in Segré, *Nuclei and particles*, 278, 1964.
Taagepera, Rein, and James L. Ray (1977), A generalized index of concentration, *Sociological Methods and Research* 5, 367-384.
Taagepera, Rein, and Steven Recchia (2002), The size of second chambers and European assemblies, *European Journal of Political Research* 41, 165-185.
Taagepera, Rein and Kaili Roopalu (2002), The number of parties and the size of the largest. Manuscript.
Taagepera, Rein, George M. Shiffler, Ronald T. Perkins and David L. Wagner (1975), Soviet-American and Israeli-Arab arms races and the Richardson model, *General Systems* 20, 151-158.
Taagepera, Rein, and Matthew S. Shugart (1989), *Seats and Votes: The Effects and Determinants of Electoral Systems.* New Haven: Yale University Press.
Taagepera, Rein, and Matthew S. Shugart (1993), Predicting the number of parties: A quantitative model of Duverger's mechanical effect, *American Political Science Review* 87, 455-464.
Taagepera, Rein, Robert S. Storey and Keith G. McNeill (1961), Breakdown strength of caesium iodide, *Nature* 190, 994.
Taagepera, Rein, and Ferd Williams (1966), Photoelectroluminescence of single crystals of manganese-activated zinc sulfide, *Journal of Applied Physics* 13, 3085-3091.
Theil, Henri (1969), The Desired Political Entropy, *American Political Science Review* 63: 521-525.
Wagner, David L., Ronald T. Perkins and Rein Taagepera (1975), Complete solution of Richardson's arms race equation, *Journal of Peace Science* 1, 159-172.

Endnotes

[1] The flow has been quite uniform. Of the 21 separate models or indices presented, the 1970s (and the late 1960s) account for 7, the 1980s for 4, the 1990s for 5, and the 2000s already for 5.

[2] Rein Taagepera, *Softening without Liberalization in the Soviet Union* (1984); Romuald Misiunas and Rein Taagepera, *The Baltic States: Years of Dependence 1940-1990* (1993); Rein Taagepera, *Estonia: Return to Independence* (1993); Rein Taagepera, *The Finno-Ugric Republics and the Russian State* (1999a).

[3] A technical difficulty is that determination of N becomes uncertain when data include a large "Other parties" bracket. However, methods to minimize the uncertainty have been devised (Taagepera 1997a).

[4] Among my other attempts at index building, elaborations on the effective threshold of representation (Taagepera 1989, 1998a,b, 2002a) are likely to have an impact. Effective magnitude (Taagepera and Shugart 1989: 126-141, Taagepera 1998a) has been used by many, but needs adjustment, so as to take into account the assembly size. A "generalized index of concentration" (Taagepera and Ray 1977) has been overtaken by the effective number. An index of imbalance was designed to discriminate in cases where inequality and concentration indices both fail to do so (Taagepera 1979a); but I have found little occasion to use it later on. "Effective size" (Taagepera and Grofman 1981) connects the effective number of components to an earlier paradox noted by Feld and Grofman (1977) and tries to distinguish between average density and crowding (bunching). Measurement of the latter remains a sore point. When a country has an ethnic minority of 20 %, it matters very much whether the minority "density" is a uniform 20% throughout the country or 90 % in some provinces and near-zero elsewhere.

[5] Hence the empirical best fit has a power exponent higher than 1/3=0.33. The best fit by Dahl and Tufte (1973: 80-82) corresponds to $F=0.25P^{0.40}$. This agrees with their logical semiquantitative hypothesis: "The size of parliaments increases with the population of a country, but at a lower rate." They did not try to answer the question why this rate (of logarithms) is around 0.4 rather than, say, 0.2 or 0.6. The communication channel model that leads to the cube root law does answer this more specific question. The simple model does not fit reality in a perfect way but is in the ball park.

[6] A few days before Christmas 1970 I read in Jean Blondel's *Comparative Politics* (1969) that the main endeavor of legislative assemblies is not passing laws but communicating. Upon that, at lunch in the UCI cafeteria I scribbled the cube root model on the back of a notice about cafeteria closing for Christmas.

[7] To what extent can $S=(nF)^{0.5}$ be said to "explain" anything in a substantive way? How does it differ from just postulating that S may depend on n and F, and then finding that R^2 is maximized by logging both variables? The difference is that here we predict not just any connection with n and F but a specific one. Empirically determined coefficients in the more general formula $S=AF^a+Bn^b+CF^cn^d +D$ may yield a somewhat higher R^2, but without any explanation of the coefficient values obtained. In contrast, here we predict ahead of time that, in the above equation, A=B=D=0 and c=d=0.5 **--** and it comes close to the actual median. This is not the entire explanation we'd like to have. Why do some countries weigh F much higher than n, and vice versa? Answering this would be the next step. But we already have explained the median value.

[8] In Figure 3 the empirical equation corresponds to $Imports/GNP=40/P^{1/3}$. The entropy model also shown offers a poorer fit and is not discussed in the present overview.

[9] The logical model is based on a visualization of exports, while testing uses imports, because formal export figures exclude the export equivalents such as labor export and tourism profits. To the extent the model applies, the characteristic absorption number varies from 160 in Uruguay to 1700 in Belgium. Using areas rather than populations of countries, the median characteristic transport distance of goods and services is found to be 58 miles (100 km), varying from 310 miles for sparsely populated Canada to 10 miles in Haiti.

[10] Actually, the rank-size expression has been an empirical rule rather than a "law" in the sense of being justified by a rational model **--** and others noticed this regularity before Zipf did.

[11] In Figure 4 the secondary cities fit the model in Cuba, Belgium and also Greece within a factor of 2, while falling markedly below prediction in Cameroon, Ivory Coast and most markedly in Madagascar. Largest cities exceed the

average primacy correction in Greece, Cuba and Ivory Coast, fit in Belgium, and fall below in Cameroon and Madagascar. One should distinguish between the absolute primacy (compared to country population) and relative primacy (compared to other cities). Madagascar is high on relative primacy but low on the absolute.

12 The reader who has published a fair number of journal articles might be interested in testing the model with her/his own vita. How close is the number of journals in which you have published to the number of articles published in the journal with most articles? How well does either figure (times its logarithm) predict the total number of articles? How well does the geometric average of the two approaches predict it?

13 The data are for linguistically Finno-Ugric peoples within the Russian Federation. Subjected by Russians at least since 1600, their external conditions have been quite similar. Core population is taken as those living in their titular area and speaking their own language as main language. Weighted circulation takes into account frequency of publication and whether it's a journal or a magazine (see Taagepera 1999a: 401). For peoples of less than 10,000 the yearly circulation levels off around 1000, because it depends mainly on outside financing.

14 When I pointed out the faster than exponential population growth in a class around 1975, a student piped up: "But if it doesn't follow the exponential pattern, what pattern *does* it follow?" Indeed. By the time I reached my office, I had the answer. The lesson: We tend to get so used to facts of life that we give up asking "Why?" when we should.

15 The bending away from the straight line since 1950 cannot be accurately mapped until future population figures become known. This is due to my using doubling times *after* a given date. But instead, one can use doubling times *before* a given date -- and this is what I'll do when returning to this issue.

16 Within the long-term exponential trend, three sudden decreases occur in N_A: around -3000 (invention of "government at a distance") , -600 (breakthrough in bureaucratic delegation of power), and +1600 (oceanic shipping). These dates introduce three successive stages. In the case of N_P, where fluctuations are wider, the 2nd and 3rd stages blend.

17 The effective number of components depends heavily on the largest components. Hence the areas of the 5 to 10 largest polities at the given time, along with total dry land area, are sufficient to determine the effective number based on areas. The procedure was repeated using estimates of populations of those polities (and the world).

18 It now occurs to me that even when the equational forms cannot distinguish between mutual and self-stimulation, the ranges of the coefficient values could discriminate between them. At the time, I did consider several refinements of Richardson's model (Taagepera 1979-1980), all with some problems.

19 The inverse square law was first presented in a conference paper (Taagepera 1987) that was supposed to be expanded in cooperation with Arend Lijphart. But we were both busy with other things, and the paper was overtaken by the need for a short overview in Taagepera and Shugart (1989).

20 There is also feedback from model to measuring. Lijphart (1999) considers two ways to measure cabinet duration. The more generous "Life I" fits the inverse square law; the more restrictive "Life II" does not. If one is interested in prediction, operationalization I is preferable to II (cf. Taagepera 2002b).

21 The model fits pretty well countries where the entire country is a single district (The Netherlands, Israel). It underestimates p' when there are many multiseat district, because resources from outside the district can help a party win a seat where it is inherently weak. A logical correction is given in Taagepera and Shugart (1993), but can be neglected when our main focus is on the nationwide number of parties.

22 If we have further information in addition to M itself, we'll of course make use of it. Thus, if we know that district-wide plurality rule is used (instead of the much more usual PR), the answer shifts to p'=1 for any M.

23 Taagepera (2002a) includes 14 systems with M>1 and 16 systems with M=1 (but different values of S). Space considerations for this article caused deletion of analysis of the $p=(MS)^{0.25}$ aspect, which remains unpublished. In all but one case the actual number is within a factor of 2 of the estimate. The actual average is high by 22%, compared to the model.

[24] Figure 16 uses the form log(L^2p)/log$p$, which is zero when $L^2p=1$. This expression becomes -1 when all p shares are equal; it approaches +1 when the largest share L approaches 1. The actual mean is around -0.006, and the actual median is +0.015, both close to the expected 0.

[25] This may not always be the optimal way to proceed, because sometimes the actual pattern is closer to linear than to exponential. One might combine the numbers of seats resulting from arithmetic and geometric approaches, adjusting their relative weights so as to make the sum of seats equal to 100%. A more elegant way might use "fractional logarithms" (Taagepera 1973b). One can take logarithms once (log x), twice (loglog x), or not at all (just x). Use notation x = $\log^0$ x, log x = $\log^1$ x, and loglog x = $\log^2$ x., so that he general expression is $\log^n$ x. Now define an operational way to determine $\log^n$ x for fractional values of *n*. For Finnish data in Figure 17 a good fit, summing up to about 100% is obtained with $\log^{0.6}$ R.

[26] I define the ideally simple electoral system to include the following (Taagepera 2001b). The *S* seats are allocated in districts of equal magnitude *M*. Ballot is categorical. Seats are fully allocated in districts, by PR. There are no legal thresholds or primaries. Parties are distinct, with no alliances or internal competition.

[27] Political scientists tend to have difficulties with imaginary situations, patiently or impatiently pointing out their unreal nature. In contrast, physicists are at ease not only with unachievable conceptual boundary cases but even with "Maxwell's demon", if it helps to visualize the situation. Incidentally, New England town meetings are assemblies consisting of all voters.

[28] The seat-vote equation emerged from my 1969 Master's thesis in international relations at the University of Delaware, in the wake of my Ph.D. in solid state physics. The SVE certainly owes to conceptual tools acquired in physics. I thank the UD Department of Political Science for graciously accepting a thesis that had little to do with international relations.

[29] The analogue of the number of seatwinning parties, the number of "vote-getting" parties, cannot be estimated.

[30] Indirectly, the suitability (if not the "truth") of *N* can be tested. It leads to good correlations with cabinet duration (Section 4.7) and issue intensity (Section 5.5), and the effective number of polities based on areas and on populations also correlate (Section 4.3). One could run the same correlations with the effective numbers based on the sum of cubes etc. If one of them should yield better correlations than *N*, then a switch should be considered.