Title
Ignorance-Based Quantitative Models and Their Practical Implications

Journal
Journal of Theoretical Politics, 11(3)

ISSN
0951-6298

Author
Taagepera, Rein

Publication Date
1999-07-01

DOI
10.1177/0951692899011003008

Peer reviewed

Journal of Theoretical Politics 11(3): 421–431
0951–6928[1999/07]11:3; 421–431; 008669

LEARNING NOTE

IGNORANCE-BASED QUANTITATIVE MODELS AND THEIR PRACTICAL IMPLICATIONS

Rein Taagepera

ABSTRACT

Some basic models in science are not grounded in what we know but in careful specifying of the limits of our ignorance. Boundary conditions on the conceptually possible range determine whether a distribution is likely to be closer to normal or lognormal and whether arithmetic or geometric means should be used. If finite and non-zero boundary values can be defined, then most actual cases tend to be around the center of the possible range. This approach makes it possible to estimate the number of parties in a national assembly within a factor of two, using only assembly size and electoral district magnitude. Deviations from ignorance-based base line are a signal to look for specifically political causes. This approach may be of use in many aspects of political science.

KEY WORDS • boundary conditions • number of parties • theoretical estimates

This topic may surprise. It might even be mistaken for an all-out attack on quantitative theoretical models as ignoring the 'qualitative nature' of politics. The purpose of building models is to express knowledge rather than ignorance, isn't it? Not quite. Sometimes we build models to make explicit the precise boundary between the known and the unknown so as to make the most of the shreds of knowledge we have. We often know more than we think we do. We abdicate important information to the realm of ignorance by considering it trivial when it is not. Careful delimitation of what we know and do not know has simple but extensive practical implications – and not only for formal modelers as such. If you have ever hesitated between arithmetic and geometric mean, or between a normal and lognormal approach, this note may offer some guidance.

My purposes are the following. First, the ignorance-based nature of some models widely used in natural and social sciences is briefly pointed out. The use of *boundary conditions* is brought in next. These are limitations on the range of outcomes. They often may look trivial but radically increase our amount of information – from nothing to something. To the extent that average conditions tend to be more prevalent than extreme ones (i.e. those close to the boundaries of possible occurrence), fruitful estimates about the outcomes can be made when one knows the possible range between the boundaries. The boundary conditions also

I thank Arend Lijphart for encouragement to publish this note.

largely determine whether one should follow an approach based on addition (arithmetic means and normal distribution) or multiplication (geometric mean and lognormal distribution).

Specific examples from the realm of political science will illustrate the general points, which apply more broadly. These examples focus on estimates of the number of parties in a country, to the extent this important number is affected by electoral rules.

This note does not pretend to break new ground but points out some features and approaches that may have been underused or misused in political science. It is written in non-technical language and is addressed to the less mathematically inclined in the profession. For those who know their mathematics, all this may be too self-evident to be even worth mentioning. And here's the rub: such insights are rarely passed on to the layperson.[1]

Some Widespread Models Based on Ignorance

Surprisingly often, basic models in science are not based on what we know but on carefully specifying what we do not know. Information means deviation from randomness. But randomness, the absence of a specific pattern, also has its rules. In statistics, the ubiquitous normal distribution is not derived from knowledge about how things like to spread themselves but on absolute lack of such knowledge. In physics, the law of ideal gases is derived from our lack of knowledge about the specific movement of molecules. In biology (and various social sciences), exponential growth is an expression of our lack of knowledge of whether the relative (percentage) growth rate increases or decreases over time.

If we do not know whether something moves left or right, our best bet might be that it remains where it is. If we do know that it moves in a certain direction with a certain speed, our best guess is that it will continue doing so (unless we have some information to the contrary) – this is the gist of Newton's First Law. If we do not know whether relative growth speeds up or slows down, our best guess is that it will remain constant: $\mathrm{d}M/\mathrm{d}t = k$. The exponential growth formula automatically results. Similarly, if we do not know about a decay rate, we assume a constant relative rate.

Not surprisingly, nature often obliges: it is as blind as we are. Distributions often are normal. Growth often is exponential. Radioactive materials do decay exponentially. When observed behavior deviates from the one expected on the basis of ignorance, we learn something new and valuable. A two-humped distribution of sizes tells us that what we thought a single population actually consists of two different ones – and we should try find out what else distinguishes them. Similarly, a non-constant growth rate causes us to ask: Why?

1. Part of this note could be couched in Bayesian terms, or reference could be made to the classic theory of measurement or to econometric approaches. But the intended readership here are those political scientists for whom such terms have little meaning and who yet at times face some of the data-processing or model-building quandaries discussed here in simplistic terms.

Thus, expectations based on ignorance supply us with a *base line* against which our observations can be compared. Without them we would lack an anchor. We could not distinguish the extraordinary from the usual, the complex from the simple, the structured from the random.

But even when the pattern comes out as expected on the basis of ignorance, we may obtain further insights. Normal distributions come with different means and standard deviations. Exponential growths and decays come with different rate constants (or doubling times or half-lives). When we determine these parameters by measurement, we can compare and contrast various growth processes, all of them exponential but with different rate constants. The model based on ignorance yields knowledge.

Boundary Conditions on Distributions: Addition versus Multiplication

Our quantitative thinking largely follows two patterns: addition and multiplication. A hesitation between arithmetic and geometric means is hesitation between additive and multiplicatory thinking. The same issue arises when one hesitates between saying that something increases by 100 percent, or by a factor of two.

Taking logarithms transforms multiplication into addition. Why are some quantities distributed according to the normal and some others the lognormal pattern? When should we graph variables on regular and when on logarithmic scale? It is a matter of addition versus multiplication – and also one of zero versus minus infinity, as will be now pointed out.

The normal distribution extends from minus to plus infinity. Its mean may have any positive or negative value, and the standard deviation may be narrow or wide, but in all cases a vanishingly small number of items will still take very large positive or negative values, in principle. If we try to fit the height distribution of people to the normal pattern, it will predict that occasionally a 10-foot person will materialize, and so would (vastly more rarely) a 100- or a 1000-foot person. This presents few problems, because the probability is so low that the universe will end before such a person materializes. Fairy tales can feature such giants. But the problem becomes much more severe toward the other end of the scale. The normal distribution also posits that, with low probability, people with negative height will materialize. Even fairy tales cannot handle that. It breaches a universal *conceptual boundary condition*, because negative sizes lack meaning. The conceptual range for size extends from zero to plus infinity, not minus to plus infinity as assumed by the normal distribution.

We can overcome the problem by noting that the logarithm of zero is minus infinity. When the conceptual range of a variable is 0 to $+\infty$, the range of its logarithm is from $-\infty$ to $+\infty$, so that this logarithm can have a normal distribution. In such a case the variable itself would follow the *lognormal distribution*, sloping gradually toward $+\infty$ (like the normal distribution) but falling more steeply to perfect zero probability when the variable itself reaches zero. The lognormal distribution asserts that zero-size people do not exist, even with a vanishingly small probability. This makes conceptual sense.

In principle, a lognormal distribution can be expected to yield a better fit than

normal distribution whenever a variable faces a conceptual lower limit at zero.[2] Then why is it that we still try to fit the heights and weights of people with normal distribution, and get away with it without visible inconsistencies? The answer is that the mean is much larger than the standard deviation. Under such a condition the issue of zero- or negative-size people becomes as remote as that of 100-foot people. In other words, lognormal and normal distributions become quite similar when the latter's standard deviation is many times smaller than the mean. Then, for simplicity, we can shift to normal distribution.

This is not true, however, for many sociopolitical variables. An attempt to fit the size distribution of countries with a normal curve may produce an apparent standard deviation larger than the mean, implying that a fair proportion of countries have a negative size. The quality of such a fit is, of course, more than poor, given that the normal curve extends into the negative region, while the data do not. However, this has not been clear to all practitioners of statistics in political science. I have seen items like telephones per capita reported with standard deviations larger than the mean. When the conceptual impossibility is pointed out to the author, the response at times is a shrug: This is what my canned program produces, and this program was prepared by competent statisticians. Well now, a screwdriver may have been produced by a highly competent craftsman, but using it as a chisel still remains poor workmanship. Misusing the normal distribution when lognormal (or possibly something else) is called for is poor scholarship – and is prone to lead to erroneous conclusions.

One could carry the reasoning a step further. There are quantities like the number of parties, for which the conceptual range extends from 1 to infinity. Given that $\log 1 = 0$, the *logarithms* of such quantities might be expected to be lognormally distributed, and one would have to go to $\log\log x$ before a normal distribution could be expected.[3]

The issue of *geometric versus arithmetic mean* is related to the previous. When normal distribution applies (at least as a good approximation), the arithmetic mean makes sense, because it reflects the center of the symmetric distribution. It is the median. This is the case when the ratio of the largest to the smallest values observed remains moderate. However, when one has to resort to lognormal fitting, the geometric mean must be used, because it reflects the center of the distribution of the logarithms (which is symmetric) and corresponds to the median. The geometric mean is always advisable when the ratio of the largest to the smallest entry is large (say, over 10) – even when the best fit deviates from lognormal. In such a situation

2. Some other distribution might be even more suitable. The point here is to explain why normal distribution makes no sense when its apparent standard deviation is larger than the mean, because it implies existence of negative values for a conceptually non-negative quantity. Fitting the logarithms of the original variable with a normal distribution (implying lognormal distribution of the original variable) overcomes this gross discrepancy. Still better fit may be obtained with some other distribution.

3. Actually, a further complexity enters. In the case of physical size, the boundary value of zero is conceptually (not to mention practically) unreachable. In contrast, for the number of parties the boundary value of 1 is reachable in principle and even in practice. I will not belabor this issue any further, except for pointing it out.

the arithmetic mean would basically depend on the largest entries, regardless of the precise size of the smallest. This is true, for instance, of the sizes of countries and of telephones per capita.

When negative values are conceptually possible the geometric mean cannot be used. The sticky point comes at zero size. Whenever even one entry is zero, the geometric mean becomes zero. A country with zero area or population will be considered non-existent and thus presents no problem. But a country could conceivably lack even a single telephone. Because of large disparities between telephone ownership in developed and underdeveloped countries, we should use the geometric mean, but this is not possible because of the single zero-telephone country! Even if we arbitrarily assign this country one token telephone, it still influences the geometric mean unduly. At the same time the arithmetic mean depends practically only on the developed countries. I have no solution short of defining an arithmetic operation intermediary between addition and multiplication (which is actually feasible). At the very least, one should be able to recognize when one has such a fuzzy situation, so as to be cautious in interpreting the results.

Finally, there is the issue of '*off by 100 percent*' versus '*off by a factor of 2*'. For small differences it does not matter. Being off by +10 percent means being off by a factor of 110/100 = 1.10, while being off by -10 percent means being off by 100/90 = 1.11; the difference is small. However, being off by +50 percent means being off by a factor of 1.50, while being off by a factor of −50 percent means being off by a factor of 2.00, which is quite a difference. Does it matter? If a country's GNP drops by 50 percent during a severe crisis and then increases again by 50 percent, has it recovered? It may look so but it has not. The GNP first dropped by a factor of 2, and then increased by a factor of 1.50, so that the GNP is only at 1.50/2.00=.75, i.e. at 75 percent of the pre-crisis GNP. Errors of this type are encountered in political science literature.

In general, if in doubt, use the multiplicative ('by a factor of . . .') approach rather than the additive. Note that an increase or decrease by a factor of k are both always possible. In contrast, a 200 percent production increase is possible, while a 200 percent decrease is conceptually impossible, because it would imply negative production. Yet a 200 percent decrease in one's wealth *is* possible: It would mean that one has fallen deeply in debt. Here the additive approach may make more sense. Once more, the main point is to recognize when one is treading on shifting sands.

As for *choice of scale* or graph paper for visualization of data, use logarithmic scale when all values are positive and the largest values are many times larger than the smallest. In other cases the regular scale is simpler – and it is inevitable when negative (or zero) values occur.

All the issues discussed in this section hinge on addition versus multiplication. Arithmetic means, percent changes, regular graph scales and normal distribution involve addition. Geometric means, changes by a factor of k, logarithmic graph scales and lognormal distribution involve a philosophy of multiplication. This is not an all-encompassing revelation but a rough aid in making practical decisions in data-processing and model-building. Next, the boundary conditions approach will be applied to estimating reasonable averages.

Median Estimates Based on Boundary Conditions

If I asked you to guess the weight of an animal called *orav* in Estonian, most readers may protest that the task is impossible because you do not know Estonian. However, your ignorance is not complete. You were told that the quantity to be guessed at is weight and that *orav* is an animal. I will make it even easier by specifying that *orav* is a mammal.

While we may not know how large something is, we often know how large or small it cannot possibly be – and this is more valuable than one may think. By establishing the bounds between the possible and the impossible one limits the conceivable range appreciably. If the distribution is close to normal, lognormal or otherwise single-peaked, most items are relatively close to median. Hence, by judiciously estimating the median within the permissible range one is likely to hit fairly close in most cases. Just knowing that something is between zero and infinity, however, will not help us. Both boundaries must be finite and non-zero.

In the present case, we know that weight is a positive number. Apart from this conceptual boundary condition, as discussed earlier, we also have the *observational boundary conditions* for mammals. The heaviest mammals are whales, who weigh up to 30 metric tons. The smallest adult mammals are shrews, who weigh about 3 g (half the weight of a regular pencil). These then are the limits, when reduced to the same unit: 0.003 kg and 30,000 kg. Which mean should we calculate?

Since negative weights are conceptually prohibited, we might expect a lognormal distribution, meaning that geometric mean is indicated. Indeed, the arithmetic mean of the possible extremes would be 15,000 kg – a fair-sized whale – regardless of whether we pick a shrew or a horse as the lowest limit. The geometric mean is around 10 kg – about 25 lb, the weight of a typical dog.

Having no other information apart from *orav* being a mammal, 10 kg is the best guess we can make. We have no grounds to argue that this particular mammal is smaller or larger than the median. We could be off by as much as a factor of 3000, if *orav* happens to be a great whale or a shrew. Actually, we are likely to be much closer. Within a factor of 10 (i.e. 1–100 kg or 2.5–250 lb), anything from a rabbit to a human would fit in. Within a factor of 100, rats and horses would be included.

Now I will let *orav* out of the bag. It means a squirrel, and most weigh less than a pound (0.4 kg). We were off by a factor of 30. Compared to pleading ignorance, this is pretty good. If you suspect I cheated by purposefully picking a mammal close to median weight, consider how many mammals would fit within a factor of 30, from squirrels to goats.

Estimates of the Largest Component

Let us now move closer to social sciences. Consider a federation like the United States, with 250 million population (in 1991) and 50 components (states). What is the likely population of the largest component?

The limits on our ignorance are the following. The mean population of the components is 5 million. This is the lowest possible size of the largest component, occurring when all components are equal. The highest possible is close to 250

million, when all other components are negligible. Since 250 is many times larger than 5 (and negative values are conceptually prohibited), the geometric mean of the boundary values is our best guess (in the absence of any other information). It is 36 million. The actual population of the most populous state (California) in 1991 was 30 million – only 17 percent lower than the estimate.

Is it luck, or can we indeed make such estimates within a 20 percent range of error? Let us first develop a general formula. By the reasoning above, the size of the largest component (L), when a total size (S) is divided into n parts, is expected to be around

$$L = [S(S/n)]^{0.5} = S/n^{0.5} \tag{1}$$

Let us now test a few other federations. In Canada, 27 million people were distributed in 1991 among 10 provinces and two very sparsely populated territories. Excluding the territories yields L = 8.5 million, while including them results in L = 7.8 million. Ontario actually had 9.7 million, an excess of 14–24 percent. For Australia (on the basis of eight components, including Canberra and the North), L = 5.95 million, while New South Wales actually had 5.73 million – a shortfall of 4 percent. On the basis of six states, L = 6.88 million, so that the shortfall becomes 17 percent.

In terms of area, the largest US state is expected to have 510,000 square miles (mi^2), and Alaska has 570,000mi^2 – high by 14 percent. For Canada the expectation is 1,100,000mi^2 for the largest component, and Northwest Territories have 1,270,000 – an excess of 15 percent. For Australia, L = 1,050,000mi^2 for n = 8 and 1,210,000mi^2 for n = 6, while Western Australia has 975,000mi^2 – a shortfall of 7–19 percent. Adding these deviations algebraically results in a mean deviation of about 1 percent, suggesting that L may indeed represent the average outcome.

This is not always the case. For the 15 union republics of the former USSR (2.24 million km^2), L = 580,000 km^2, while actually the Russian RSFR had three times this area: 1,708,000 km^2. The same applies to Prussia within Imperial Germany. This is the analogue of finding a two-humped distribution of heights in a population. In the latter case the non-normal distribution tells us that two distinct populations are involved (possibly female and male). In the Soviet and German cases, the deviation from the ignorance-based estimate tells us that something else enters, beyond randomness. Indeed, both entities involved a large imperial core that gave formal autonomy to its latest acquisitions (the ones that maintained some historical or ethnic distinctiveness), without dividing up the core itself (Prussia and the ethnically Russian area, respectively).

And this is the point. Anyone can look up Alaska's exact area in an almanac, instead of calculating an estimate. But now we have a *base line* to distinguish between different ways of subdividing an entity. By pinning down what can be inferred on the basis of ignorance, we are able to specify the extent of new information.

For our purposes, this simple model is also a preparatory step for a more complex one, where the boundaries on ignorance are invoked repeatedly and the topic is of central interest in political science: the number of parties. Indeed, if one had to give a single number to characterize the politics of any country that employs competitive elections, it would be the number of parties active in its national assembly.

The Number of Parties

In a national assembly of S seats, how many parties (and independents) would one expect to be represented with at least one seat? The first reaction might be that this is impossible to say without knowing the country. Still, we know already that this number (n) must be at least 1 and at most S. In the absence of any other knowledge, we might try the geometric mean, which is $S^{0.5}$, but it yields overly high values in most cases. With approximately 400 seats, the US House does not have 20 parties or independents. Further information is visibly needed on what is inside the black box.

Most countries divide their territory into electoral districts. This is where the allocation of seats takes place, not nationwide. Consider the simple case where all seats are allocated in districts of equal magnitude (M, the number of seats allocated within the district), with no legal thresholds or other complications. In such a district, the boundary conditions are the following: one party winning all M seats, and M parties winning one seat each. The conceptually allowed range is $1 \leq n \leq M$. Since n cannot be negative and, moreover, M in multi-seat districts can be appreciably larger than one, geometric mean would be the best guess: $M^{0.5}$. For single-member districts this is true by definition: $n = M = 1$. The largest M occurs in the Netherlands where the entire country is a single district of $M = 150$ (with a modest legal threshold), and indeed, close to $150^{0.5} = 12$ parties and independents tend to win seats. Countries with intermediary district magnitudes also agree approximately. But we are interested in the national assembly, not the single district.

Therefore, the reasoning is repeated. The assembly is expected to have at least as many parties as a single district. As the upper boundary, the country is not expected to harbor more parties than it would if the entire country were made a single district, Netherlands-style. Hence, on the nationwide level, the expected range is $M^{0.5} \leq n \leq S^{0.5}$ and the geometric mean is

$$n = (MS)^{0.25} \tag{2}$$

This result was first proposed in Taagepera and Shugart (1993). A more recent study (Taagepera, 1998) of 14 regimes with multi-member districts and 16 with single-member districts yields values that range from one-half to double the estimated value in 28 out of the 30 cases. For the $M > 1$ group the actual numbers of seat-winning parties surpass the estimates by an average of 18 percent: their arithmetic mean is 5.78 as compared to the estimated mean of 6.81. For the $M = 1$ group the excess is 26 percent (4.88 versus the expected 3.88).

Is this sufficiently close? It depends on the purpose. As with Alaska's area, if we want the precise value for the given *past* election, we can look it up. But we cannot look it up for *future* elections. If we want to draw general conclusions regarding the impact of assembly size and district magnitude, then the present estimate has made considerable progress, compared to 'anything between 1 and S is possible'. And we have achieved it with remarkably little information – just assembly size and district magnitude – by squeezing the maximum out of this information.

Indeed, no more precise predictions should be expected, unless one is a fanatical institutionalist. Country-specific factors such as sociopolitical and geographic

heterogeneity, culture and history should be expected to have an impact. Other institutional factors also enter, such as parliamentary versus presidential regimes and varying allocation rules used in one- or two-round elections. This being so, it is truly remarkable that the actual values do not fluctuate by more than a factor of two around the estimate based on S and M alone!

Now the second role of the ignorance-based estimate enters – supplying a base line. The discrepancy between the estimate and Russia's actual size within the USSR makes us ask 'Why?' The same question arises in the case of countries with unusually many or few parties. There is little to wonder about countries where the actual number pretty much agrees with the estimate, such as Portugal 1975–87 (6.9 versus 7.3) or Canada 1878–1988 (4.4 versus 4.0). But why did Imperial Germany (to take the most deviant case) have an average of 13.6 seat-winning parties rather than the estimated 4.5? In the other direction, why did the US House always have few parties and independents, starting with an average of 3.0 (versus an estimated 3.9) in 1828–82 and dropping to 2.5 (versus estimated 4.6) in 1938–88? When the average effect of assembly size and district magnitude is removed, the residues highlight the impact of the other factors, making it easier to unscramble them.

One sometimes encounters the objection that the estimates based on boundary conditions are blindly mechanical so that, even if they produce results, they do not 'explain' the outcome in terms of specifically political mechanisms. Yet the same people accept a normal distribution of some political variable without requiring a political explanation. Indeed, there is nothing to explain. Only a deviation from the normal base line would be begging for a political (or other more specific) explanation. The same is true for the boundary conditions. Being far away from boundaries is to be expected. Being close to them would require explanation. Such an approach was actually used here. We first tried the geometric mean of boundary conditions 1 and S, but it did not work. Then we introduced a minimal amount of specifically political knowledge: the notion of electoral districts. It brought us into the right ballpark. Beyond this base line, more specific explanations have to take over.

Taagepera and Shugart (1993) went beyond the number of seat-winning parties to estimate the 'effective number of parties' (N) that has received wide currency (Lijphart, 1994: 70; Cox, 1997: 29). The effective number is smaller than the number of all parties that win at least one seat, because a procedure of self-weighting emphasizes the role of the larger parties: $N = 1/\Sigma s_i^2$, where s_i is the ith party's fractional seat share. The procedure for estimating N is outlined here only briefly. It applies the ignorance-based approach in two further stages. From the previous estimate of the number of seat-winning parties, $n = (MS)^{0.25}$, the number of seats going to the largest party is estimated, using the previous equation $L = S/n^{0.5}$. In turn, L sets boundary conditions for the effective number of parties. Finally, the estimate for the effective number of parties emerges as

$$N = (MS)^{3.16} \tag{3}$$

Supplemented with a 0.85 multiplier to account for the competition between the two top parties, this expression reflects the average N for a number of regimes (Taagepera and Shugart, 1993: Figure 2).

Graphical Representation of Boundary Conditions

It is often useful to present data visually, y versus x. Patterns may hit the eye that no inspection of correlation and regression coefficients could detect. In the case of such a graphical display, it might be useful to include the conceptual boundaries, shading off the *forbidden regions*. Sometimes the information thus added to the graph may look trivial, such as showing that percentages below 0 and above 100 are forbidden. But sometimes the explicit showing of boundaries elicits new ideas for modeling – and it avoids mistakes like extending a regression line into obviously forbidden regions.

The same purpose is served by conceptual *anchor points*, by which I mean points that any data fit logically must include. For example, in any democratically elected assembly, a party that obtains zero votes should get zero seats. Conversely, if a party should get 100 percent of the votes, it should also win 100 percent of the seats. In the intermediary range such proportionality usually does not apply, because practically all electoral systems give a bonus to the larger parties. Thus a regression curve based on actual data points may not pass through the anchor points (0;0 and 100;100), unless this is stipulated as a requirement. And it should be stipulated, because we know that, unfailingly, any party with zero votes will get zero seats (unless there is dictatorial monkey business!) and such knowledge should not be tossed out.

Conclusion

My focus has been on one political science issue where the payoff of the ignorance-based approach is clearest: the number of parties. This approach is likely to prove useful in various other aspects of political science. In physics, this approach sometimes leads to impressive results like the following. Start with certain requirements of continuity in an electromagnetic field (Maxwell's equations). Then the stipulation of boundary conditions of a space (meaning, which parts of the wall are conducting and which are isolating) uniquely determines the field strength at any point enclosed. In social sciences a faint analogue is the simple logistic equation: stipulate a minimum (usually zero) and a maximum size, plus the size at any two points in time; then continuity considerations determine the size at any time.

We began by pointing out that some of the most usual models are based on judicious delimitation of the bounds of our ignorance, rather than any specific knowledge. When non-positive values are conceptually prohibited, the use of geometric means emerges as preferable to arithmetic mean. We then proceeded to show that geometric means of widely disparate conceptual boundary conditions often represent fair estimates of the values actually occurring. This procedure built up from trivial-looking estimates of weights of mammals and areas of federal provinces to estimates of the number of parties, which give approximate answers to important questions of how institutions affect party systems. Indeed, using only assembly size and district magnitude, much of the variation observed is removed, enabling one to have a better grip on the effect of the various other cultural and

institutional factors. Without going into detail, the usefulness of graphical representation of boundary conditions and anchor points has also been pointed out.

REFERENCES

Cox, Gary W. (1997) *Making Votes Count: Strategic Coordination in the World's Electoral Systems*. Cambridge: Cambridge University Press.

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

Taagepera, Rein (1998) 'Nationwide Threshold of Representation', unpublished.

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: 455–64.

REIN TAAGEPERA is Professor Emeritus of Political Science at the University of California, Irvine, where he has taught since 1970, and Tartu University, Estonia, where he has taught since 1992. His PhD was in physics, and his main research interest is in electoral systems. His books include *Seats and Votes: The Effects and Determinants of Electoral Systems* (Yale University Press, 1989) and *The Baltic States: Years of Dependence, 1940–1990* (California University Press and Hurst, 1993). He received 23 percent of the vote in the Estonian presidential elections of 1992. ADDRESSES: (Sept.–March) School of Social Sciences, University of California, Irvine, CA 92697, USA; (April–Aug.) Department of Political Science, Tartu University, 50090 Tartu, Estonia.
[email: rtaagepe@uci.edu and mtaagepe@ut.ee]

Paper submitted 20 March 1998; accepted for publication 2 October 1998.