

# BEHIND THE CUBE RULE: IMPLICATIONS OF AND EVIDENCE AGAINST A FRACTAL ELECTORAL GEOGRAPHY.

*John Maloney<sup>\*</sup>, Bernard Pearson<sup>\*\*</sup>, and Andrew Pickering<sup>\*\*\*</sup>*

---

\* Reader in Political Economy, University of Exeter and *Corresponding Author*. School of Business and Economics, University of Exeter, Streatham Court, Rennes Drive, Exeter EX4 4PU. Fax 00-44-(0)1392-263242. Email [J.Maloney@ex.ac.uk](mailto:J.Maloney@ex.ac.uk).

\*\* Lecturer in Statistics and Econometrics, University of Exeter. School of Business and Economics, University of Exeter, Streatham Court, Rennes Drive, Exeter EX4 4PU.

\*\*\* Lecturer in Economics, University of Bristol. Department of Economics, University of Bristol, Alfred Marshall Building, Woodlands Road, Bristol BS8 1TN.

# **BEHIND THE CUBE RULE: IMPLICATIONS OF AND EVIDENCE AGAINST A FRACTAL ELECTORAL GEOGRAPHY.**

## **ABSTRACT**

In 1909 Parker Smith showed that the ratio of seats won by the two major parties in Britain was close to the cube of the ratio of their votes. Taagepera and Shugart argue, wrongly, that a fractal electoral map implies this. In fact their premises imply that the seats' ratio will be the votes' ratio to the power of  $\sqrt{3}$ , not 3. However, in the six countries we examine, the figure is between 2 and 3. This implies that the electoral map is nonfractal, political allegiances becoming less 'clustered' as you move from a macro to a micro scale. Taking the U.K., we ask if this is due to the geographical pattern of income distribution, and find that this is even further away from fractality than is voting. This fits the well-known 'chameleon effect' whereby poor (rich) people in rich (poor) constituencies vote as if richer (poorer) than they really are.

**Keywords:** fractal, election, voting, cubic

If the UK ever has a referendum on electoral reform, the cube rule is unlikely to be on many posters or car bumpers. Quite apart from being too arcane for any politician to be able to mention it without journalistic ridicule, its main victim these last hundred years has been the runner-up in each general election. And the two main parties of the day have been happy enough with a system which has merely thrust them into a smaller minority when they have been in a minority anyway, in exchange for untrammelled power when they top the poll, even when getting considerably less than half the vote. Meanwhile, the real victims of the first-past-the-post system, the Liberals and most of the minor parties, would have given a lot to have been on the wrong end of nothing worse than the cube rule.

It was, however, in the context of electoral reform that the cube rule emerged. The idea that, when the ratio of votes gained by the two main parties was  $a:b$ , they could expect the ratio of seats to be roughly  $a^3 : b^3$ , was first put forward by James Parker Smith in his evidence to the Royal Commission of Systems of Elections of 1909. Forty years later Kendall and Stuart<sup>1</sup> put a theoretical basis of sorts under it, by showing that (a) if constituencies were normally distributed in terms of support for each of the two main parties, there existed a standard deviation that would almost precisely yield a cube rule (b) that the constituencies had had almost exactly this critical deviation at the general elections of 1935, 1945 and 1950. But what does ‘normally distributed’ mean here? Kendall and Stuart invited their readers to imagine the constituencies arranged along a horizontal line from safest Conservative to safest Labour, and a cursor moving over them as national voting

intentions changed. The number of constituencies falling to the other side with successive 1% swings in one direction would have a frequency distribution that was normal. However Kendall and Stuart gave no reason *why* the set of constituencies should have a standard deviation that gave a virtual cube – as opposed to square, quartic or anything else – rule.

Gudgin and Taylor<sup>2</sup> add an extra dimension. They use a Markov system in which your probability of voting for X is a function of whether your neighbour votes for X. The “cubeness” of the system will depend on (a) this probability (b) the size of the constituencies. After a detailed look at a constituency in Newcastle, Gudgin and Taylor conclude that this seat possesses the characteristics which, if replicated nationally, would yield something close to a cube rule. Like Kendall and Stuart, however, their concern is not to suggest reasons *why* a typical seat should be like this.

But in the 1970s and 80s Taagepera tackled the question from a new angle and did come up with a “first principles” explanation, albeit a highly tentative one, for the cube rule. In section 1 we outline this explanation, which relies on the electoral map being fractal, and suggest that Taagepera’s premises in fact lead to a “square root of three rule” rather than a cube rule. Section 2 pitches this proposition at the evidence, and section 3 interprets the results: so far as winners do better than they would under a ‘square root of three’ rule, this suggests that electoral geography is nonfractal, with political heterogeneity increasing as you move from a macro to a

micro scale. We also ask how far the geographical pattern of income distribution might explain this.

## 1. Theory

Taagepera and Shugart's *Seats and Votes* (1989)<sup>3</sup> contains a proof of the cube rule in 3 stages:

- (1) If, for all pairs of parties, the ratio of seats won is to be a consistent function of ratio of votes won, the ratio of seats must equal the ratio of votes to the power of n.
- (2) Under a first-past-the-post system, with single-member constituencies, the coefficient n will be equal to  $\ln(V)/\ln(S)$  where V and S are the total numbers of votes and seats nationally.
- (3)  $\ln(V)/\ln(S)$  will be very close to 3.

Let us take these stages one by one.

- (1) Here Taagepera draws on a proof by Theil (1969)<sup>4</sup>. Theil shows that for the ratio of seats between *all* pairs of parties in a multi-party election to bear a consistent relationship with the ratio of their respective votes, it must take the form expressed above. Let  $S_i/S_j = f(V_i/V_j)$  for all i, j, where S is a party's share of the seats and

V its share of the votes. Then, if we take three parties, A, B and C, we get:

$$S_C/S_A = (S_C/S_B) \cdot (S_B/S_A)$$

and hence

$$f(V_C/V_A) = f(V_C/V_B) \cdot f(V_B/V_A).$$

This relation will hold when and only when  $f(V_i/V_j)$  takes the form  $(V_i/V_j)^n$ .

- (2) The logarithmic proposition is Taagepera's own. Clearly, Taagepera has to make a specific assumption as to how 'clustered' party support is on the electoral map in order to reach any such conclusion. To take an extreme case, if there were no clustering, in the sense of constituencies being identical, then an infinitesimal swing in the national vote could lead to a 100% swing in seats. At the other extreme, we could imagine voters being completely clustered in the sense of voters behaving identically in each constituency, but the constituencies themselves being differentiated. This would be equivalent to each constituency containing only one voter, and first-past-the-post would produce proportional representation.

So what degree of clustering would produce Taagepera's logarithmic rule, and how does he justify it?

Taagepera looks for plausible regularities which might help us find a specific functional form for  $n$ . He says that 'in the absence of any information to the contrary, the degree of clustering of like-minded voters must be assumed to be the same on all levels (that is, a magnified piece of a political adherence map would look like the original map)'. He is, thus, assuming that the electoral map is

fractal<sup>5</sup>. We now show that such an assumption does not lead, as Taagepera claims, to the rule  $n = \ln V / \ln S$ , but instead implies that  $n = \sqrt{\ln V / \ln S}$ .

How do we give content to the idea of a fractal political adherence map? Suppose a country consists of 100 constituencies, each consisting of 100 towns, each consisting of 100 voters. A fractal map would replicate frequency distributions at each level of disaggregation. Thus if, e.g. 40% of constituencies (currently) had a Labour vote of more than 55%, so would 40% of towns in the median constituency<sup>6</sup>.

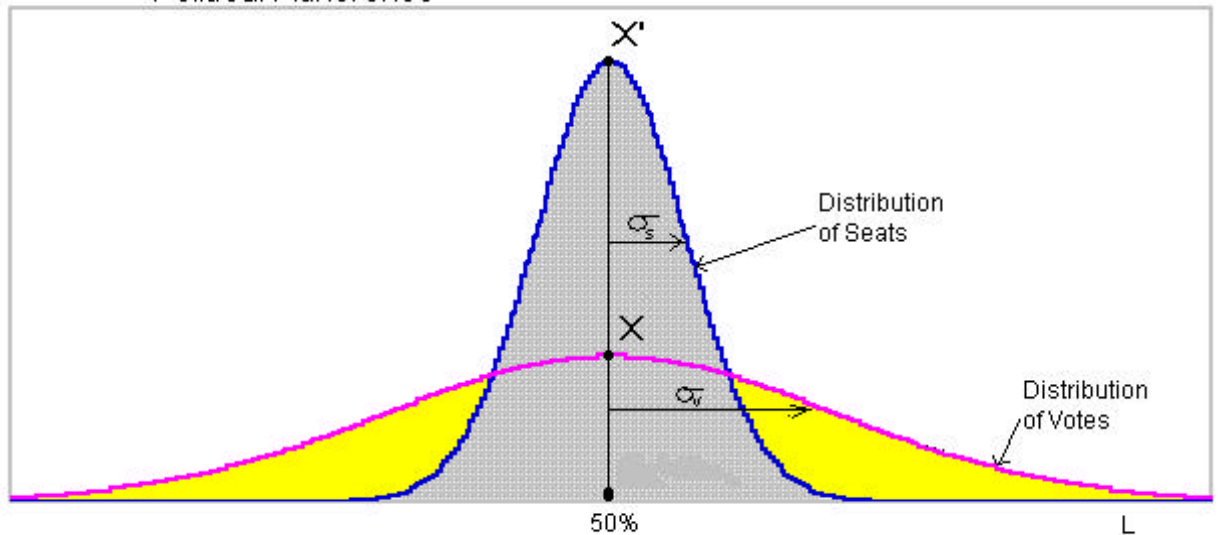
Now let  $L$  be the percentage Labour vote, and let  $\text{var}(L_1)$  and  $\text{var}(L_2)$  be, respectively, the variance of  $L$  between seats, and the variance of  $L$  between towns in any one seat. If the map is fractal,  $\text{var}(L_1)$  will equal  $\text{var}(L_2)$  and the variance of the towns nationally will be  $2\text{var}(L_1)$ <sup>7</sup>.

But what about the variance of the voters in a given town (call this  $\text{var}(L_3)$ )? Does it make sense to think of individual voters as 40% or 80% Labour? The answer is that we don't need to: all the fractal principle requires is that there be as many knife-edge (right on the margin between Labour and Tory) voters in a knife-edge town as there are knife-edge towns in a knife-edge constituency (and knife-edge constituencies in the country).  $\text{Var}(L_3)$  must therefore be represented as equal to  $\text{var}(L_2)$  and  $\text{var}(L_1)$ . The variance of voters' 'percentage attachment' to Labour, nationally, is thus  $3\text{var}(L_1)$  (see, again, footnote 7), though, to repeat, we're only

interested in what number and distribution, between seats, of marginal voters this gives us. We don't have literally to describe anyone as 57.2% Conservative<sup>8</sup>.

Suppose, then, that we scale the country up by a factor of  $y$  at each step, and carry on doing this until we are at the scale of the individual voters. If there are  $V$  voters, we will have gone through  $\log_y V$  steps, and if  $\text{var}_1$  is the variance we get when we disaggregate one stage, the variance of  $V$  will be  $\text{var}_1(\log_y V)$ . (See, again, footnote 7.) On the same reasoning, if there are  $S$  seats, the variance of  $S$  will be  $\text{var}_1(\log_y S)$ . Hence  $\text{var}(V)/\text{var}(s) = \ln V/\ln S$ .

Figure 1. Density Functions of Individual Voter and Constituency Average Political Adherence



So take Fig.1. The density function of the constituency percentage Labour votes is normally distributed with a standard deviation  $\sigma_s$ . We initially assume that its



mean is 50% (the parties are level-pegging). The density function of individual voter political adherence is distributed normally with a standard deviation  $\mathbf{s}_V = \mathbf{s}_S \sqrt{\ln V / \ln S}$ . Now assume an infinitesimal swing from Tory to Labour. Labour will pick up the fraction OX of the total vote and O'X' of the total seats (fig.1). Thus, if  $S_A$  and  $V_A$  are the shares of the vote and the seats held by Labour (and using B for the corresponding Conservative shares), then

$$dS_A / dV_A = OX' / OX$$

But  $dS_A / dV_A = n$  (given that  $S_A / S_B = (V_A / V_B)^n$ ),

and  $OX' / OX = \sigma_V / \sigma_S$ .

Hence  $n = \mathbf{s}_V / \mathbf{s}_S = \sqrt{\ln V / \ln S}$ , and not, as Taagepera claims,  $\ln V / \ln S$ .

- (3) We now proceed to the final part of Taagepera's justification of the cube rule, his proposition that, under certain reasonable assumptions, constituencies will be chosen of a size that will make  $\ln V / \ln S$  approximately 3.

Assume the system is (consciously or otherwise) designed to minimise the workload of an MP. Assume this is done by minimising the number of channels through which an MP has to communicate (i.e. each potential channel is equally time-consuming). Suppose there are V voters and S MP's. Then each MP has V/S constituents & thus 2V/S channels (one channel each way with each constituent).

But the MP also has a channel to every other MP, and furthermore must monitor all channels *between* other MP's. This gives him  $S(S-1)/2$  more channels  $\cong S^2/2$ .

The objective, then is to choose  $S$  to minimise  $C = S^2/2 + 2V/S$ . Differentiating gives  $dC/dS = S - 2V/S^2$  which is zero at  $2V = S^3$  i.e.  $\ln V/\ln S = \text{approx. } 3$ .

This third proposition is hardly overwhelmingly persuasive, and Taagepera is fairly tentative at advancing it himself *as a theory*. However, as he rightly points out, it fits a number of countries remarkably accurately in practice (more on this below.)

If, then, we accept it, and substitute the correct rule for  $n$  ( $\sqrt{\ln V/\ln S}$ ) for Taagepera's incorrect  $\ln V/\ln S$ , we have, not a cube rule, but a 'square root of 3 rule.' If the electoral map really is fractal, and we accept the final part of Taagepera's proof, a ratio  $a:b$  of votes will be amplified into a ratio  $a^{\sqrt{3}} : b^{\sqrt{3}}$  in terms of seats.

We now look at some empirical evidence. As we have said, if the political adherence map is fractal, we would get the result  $n = \sqrt{\ln V/\ln S}$ . In the tables below, we use actual values of  $V$  and  $S$ , i.e. the actual numbers of voters and MPs in each country. Because most of the countries we examine are close to the  $2V = S^3$  formula, a fractal electoral geography would imply something close to the  $\sqrt{3}$  rule; or, to put it the other way round, the empirical existence of a  $\sqrt{3}$  rule or something close to it would imply a near-fractal electoral geography.

## 2. Evidence

We first present actual data in terms of votes cast and available seats for 6 countries where non-PR constituency systems have prevailed and the political system has been dominated by two major parties. The countries are Australia, Canada, France, New Zealand, the UK and the US. Using these data we impute a value for  $n$  that would hold given fractality. We then go on to estimate actual values of  $n$  and compare the estimates with the imputed values. The data constitutes votes attained and seats won by the major parties of these countries in elections since 1955 and was obtained from Keesing's Contemporary Archives.

The data presented in table 1 refers to data for recent elections in each country. The imputed value of  $n$  is  $\sqrt{\ln V / \ln S}$ , the value  $n$  *would* take, given the actual ratio of seats to votes, if the political allegiance map were fractal. Fractality plus Taagepera's model of paranoid MPs predicts a value of  $n$  close to  $\sqrt{3}$  (1.732), and the ratio of seats to votes is close in all 6 cases<sup>10</sup> to Taagepera's prediction.

**Table 1 Voters, Constituencies and the computed value of the power coefficient,  $n$  under fractality.**

Country (Election)	# Voters (V)	# Constituencies (S)	"n"
Australia (1998) <sup>11</sup>	11547325	148	1.804
Canada (1997)	13174698	301	1.695
France (1997)	28037180	577	1.642
New Zealand (1993) <sup>12</sup>	2085704	120	1.743
UK (1997)	31288108	658	1.631
US (1996)	96400634	51 <sup>13</sup>	2.168

We now focus on the issue of fractality itself.

Our econometric test of Taagepera's model is as follows. If

$$S_A/S_B = (V_A/V_B)^n$$

then 
$$\ln\left(\frac{S_A}{S_B}\right) = n \ln\left(\frac{V_A}{V_B}\right).$$

We estimate 
$$\ln\left(\frac{S_A}{S_B}\right) = \hat{\mathbf{a}} + \hat{\mathbf{b}} \ln\left(\frac{V_A}{V_B}\right) \quad (1)$$

where  $\hat{\mathbf{a}}$  is an estimable measure of bias, and  $\hat{\mathbf{b}}$  is an estimable exponential coefficient. In the above model  $\alpha = 0$  and  $\beta = n$ . Equation (1) requires nothing more than step one of Taagepera's argument and does not require fractality. Inference concerning fractality follows from our estimates of  $\beta$ .

**Table 2. Econometric Output**

Country	Estimated Equation <sup>14</sup>	$\bar{R}^2$	obs.	DW <sup>15</sup>	Fractal n	k
<b>Australia</b> (1955-1980) <sup>16</sup>	$\ln\left(\frac{S_A}{S_B}\right) = -0.303 + 2.62 \ln\left(\frac{V_A}{V_B}\right)$ (0.023) (0.178)	0.96	11	2.05	1.80	0.78
(1983-1998) <sup>17</sup>	$\ln\left(\frac{S_A}{S_B}\right) = 0.068 + 2.64 \ln\left(\frac{V_A}{V_B}\right)$ (0.097) (1.01)	0.49	7	1.03	1.80	0.78
<b>Canada</b> (1957-1997)	$\ln\left(\frac{S_A}{S_B}\right) = 0.183 + 2.703 \ln\left(\frac{V_A}{V_B}\right)$ (0.051) (0.152)	0.96	13 <sup>18</sup>	0.81	1.70	0.74
<b>France</b> (1962-1997) <sup>19</sup>	$\ln\left(\frac{S_A}{S_B}\right) = 0.388 + 2.732 \ln\left(\frac{V_A}{V_B}\right)$ (0.236) (0.475)	0.80	9	2.40	1.64	0.72
<b>New Zealand</b> <sup>20</sup> (1960-1992)	$\ln\left(\frac{S_A}{S_B}\right) = -0.032 + 2.798 \ln\left(\frac{V_A}{V_B}\right)$ (0.041) (0.302)	0.89	11	1.66	1.74	0.71
<b>UK</b> (1955-1974(4))	$\ln\left(\frac{S_A}{S_B}\right) = -0.003 + 2.46 \ln\left(\frac{V_A}{V_B}\right)$ (0.024) (0.28)	0.93	7	1.47	1.63	0.76
(1979-1997)	$\ln\left(\frac{S_A}{S_B}\right) = 0.19 + 2.088 \ln\left(\frac{V_A}{V_B}\right)$ (0.024) (0.091)	0.99	5	2.73	1.63	0.84
<b>US</b> (1960-1996)	$\ln\left(\frac{S_A}{S_B}\right) = 0.086 + 3.65 \ln\left(\frac{V_A}{V_B}\right)$ (0.64) (2.47)	0.12	10	1.75	2.17	0.78

In table 2 we present econometric output of estimation of (1). The output relates to regressions ran upon the general election data for the two main parties of each of the six countries in our sample over the period detailed. Party A respectively refers to the Australian Labour Party (ALP), the Canadian Conservative Party, the French

right-wing ‘coalition’ of the day <sup>21</sup>, the New Zealand Labour Party, the UK Labour Party and the US Democrat Party. Party B respectively refers to the Australian Liberal and National Party (ALNP), the Canadian Liberal Party, the French Socialist Party, the New Zealand National Party, the UK Conservative Party and the US Republican Party. The sample sizes (obs in table 2) are quite small and hence the results should be treated with some caution. Elections only occur once every few years, and are regularly subject to reform, which econometrically translates as structural breaks (and found to be statistically significant in Australia and the UK). These two factors hamper the econometrics although in most cases a reasonable degree of explanatory power as measured by the adjusted  $R^2$  was found.

In the case of Australia we found that the ALP was severely disadvantaged by the electoral system prior to 1983. In particular, the constant term implies that an ALP vote was 85.4% as powerful as an ALNP vote at that time. Bias in favour of one party or another was also found in Canada (the Conservatives), France (the Socialists) and the UK post 1979 (the Labour Party). These advantages are manifested from regionalism (as in the UK and France, where the favoured party’s support has been less uniformly spread across the regions than its rival’s) or strength in smaller rural constituencies (as in Australia and Canada).

We found in general that the ‘cube rule’ could more accurately be described as a ‘somewhere between a square and a cube rule’ in all cases except the US. The econometric output of the US yielded much lower levels of significance, and

neither parameter was found to be significantly different from zero. Examination of the US data confirms the erratic relationship between votes cast and electoral college votes gained.

### 3. Interpretation

With the exceptions of the US and Australia (1983-1998) all the cases in table 2 have a statistically significantly higher value of  $n$  than a fractal electoral map would give (the values for  $n$  reported in table 1). In other words, in each case  $\sigma_V/\sigma_S$  (and hence *actual*  $n$ ) exceeds  $\sqrt{\ln V/\ln S}$ . This implies that, as you take successively smaller bits of the electoral map and blow them up to the size of the whole thing, the degree of clustering diminishes. Voters within a constituency are more diffuse in their political orientations than are constituencies within the nation.

The statistic  $k$ , in the final column, is a measure of this effect.  $k$  in each case is imputed from  $n$ . It represents the departure from fractality necessary to give  $n$  its actual value in the country concerned. So how are we quantifying departures from fractality? Suppose, again, that a country were divided into 100 constituencies, each of which were divided into 100 towns.  $k$  is then defined as the ratio of constituencies' standard deviation nation-wide to towns' standard deviation within a constituency; or, more generally, the ratio of the respective standard deviations within two pieces of the electoral map that differ by a scale of  $100^{22}$ .  $k=1$  is the fractal case.  $k<1$ , which is our result in all the countries we have looked at, implies

that the more you use the magnifying glass, the more heterogeneous the picture you see. The relationship between  $n$  and  $k$  is an inverse one <sup>23</sup>, and as  $k$  tends to infinity  $n$  tends towards unity (proportional representation). This is simply a restatement of the proposition (see above, p. 6) that zero dispersion of voter behaviour in each constituency ( $\sigma_v=0$ ) is equivalent to each constituency containing only one voter, which obviously implies PR.

In the case of the UK,  $n$  fell from 2.46 to 2.088 (and  $k$  rose from 0.76 to 0.84) between the two sub-periods. As long ago as Mr Heath's win in 1970, commentators were pointing out that the cube rule seemed to be turning into a square rule. The number of marginal seats was declining, so that a larger swing in the total vote was needed to achieve any given net gain in seats by Labour or the Conservatives (though Norris and Crewe find that the total *gross* change in seats between *all* parties did not decline in relation to percentage swing). <sup>24</sup> Later this began to be linked to a regional polarisation of political allegiances. The high point was reached in the general election of 1987, when Labour won only three seats south of a line from the Wash to the Bristol Channel (London excepted) and the Conservatives, despite an 12-point national lead in the vote, only won 10 out of a possible 72 seats north of Hadrian's wall. In 1992 and 1997 there was only a slight reversal of this polarisation - masked, in 1997, by the spectacularly unequal overall result. As the political dispersion of constituencies moves upwards to rival the political dispersion of voters within constituencies, Britain's 'over-fractal' electoral geography appears to be moving closer to 'strict' fractality. <sup>25</sup>



We have therefore estimated  $n$  for different countries and thence inferred how, and how far, their political adherence maps are non-fractal. But what factors actually determine how far away from fractality the political allegiance map is? If we are dealing with two parties whose support is closely related to voters' incomes, then the fractality or otherwise of income distribution is going to be a major determining factor.

We have only looked at income distribution in the UK so far, but we found it to be much more 'super-fractal' even than the distribution of political allegiance, the standard deviation of (log) income in the average constituency being 3.87 times the standard deviation of average constituency incomes around the national mean<sup>26</sup>. Hence, *if* people voted entirely according to their income (i.e. all those on £40,000 a year started voting Labour when Labour's lead reached 10% etc.) first-past-the-post would yield, not a cube rule, but something close to a quartic rule.

However, it is no surprise that election results in Britain are much closer to proportionality than this. One of the best-documented electoral facts is the 'chameleon effect' whereby voters partially assume the colouring of the area in which they live. People with £60,000 a year in Jarrow are more prone to vote Labour than people with £60,000 a year in Reigate<sup>27</sup>. We might thus credit our man in Jarrow (Reigate) with a lower (higher) 'effective income' than his actual one. The result will be that Jarrow's (Reigate's) *average* effective income will be

even lower (higher) than its average actual income. To this extent  $\sigma_S$  is increased by the chameleon effect, and  $n$  is reduced.

But now suppose we have a party whose support is in no way income based – its support courses evenly downwards from millionaires through meths drinkers to university lecturers. There will, nonetheless, be some characteristic, or vector of characteristics, of which the probability of an individual supporting that party will be a monotonic function. The problem is that, unless this vector is identifiable and measurable in individual cases, we will be able to make no predictions as to whether a cube rule, square rule or whatever is likely. We will always be on the firmest ground when weighing up the cube rule for predominantly income-based parties. And in many democracies, the two most unequivocally income-based parties are the two largest parties. This is probably the real reason why the ‘ $n$  rule’ has generally worked best, and has certainly received nearly all the attention, when the context is the two main parties of state.

#### **4. Conclusion**

While we contest Taagepera and Shugart’s contention that their assumptions lead to a justification of the cube rule, their ‘fractal’ yardstick against which to measure the electoral system is an original and very fertile idea. No one would expect the electoral landscape to be neatly fractal or even neatly non-fractal. But the way that the ‘attrition-of-minorities’ effect is consistently stronger than it would be in a

fractal world is significant. So are the facts that the UK, at least, is much closer to fractality than it used to be, and has always been much closer to fractality than if people had voted purely according to their incomes. Overall it seems that political diversity is greater within than between constituencies, but that this gap is narrowed by the 'chameleon effect' whereby your neighbours affect your own vote and, even after this effect is allowed for, shows signs of narrowing still further.

## FOOTNOTES

1 M.G. Kendall and A. Stuart, 'The law of cubic proportion in Election results', *British Journal of Sociology*, 1 (1950), 183-196.

2 G.Gudgin and P. J .Taylor, *Seats, Votes and the Spatial Organisation of Elections* (London, Pion, 1979).

3 R.Taagepera, and M. S. Shugart, *Seats and votes: the effects and determinants of electoral systems*. (New Haven, Conn.: Yale University Press, 1989)

4 H.Theil, 'The Desired Political Entropy', *American Political Science Review*, 63 (1969), 521-525

5 A fractal shape is made up of parts which are similar in structure to the shape itself. Each of those parts is in turn made up of similar parts on a smaller scale, and so on into an infinite regress.

6 At first sight, the concept of fractality seems ambiguous. Suppose 40% of towns in the median seat have 55% or more Labour voters. How do we replicate this one stage up the scale? Does fractality require that (a) 40% of seats have 55% or more Labour voters, or (b) 40% of seats have 55% or more Labour towns (i.e. towns with a more than 50% Labour vote)? The two will not normally come to the same thing. But in fact, only (a) is consistent with fractality, for the following reason. The number of knife-edge towns in

the median constituency must equal the number of knife-edge constituencies, or the slightest shift of national opinion will destroy fractality. Therefore the standard deviation of L (percentage Labour vote) must be the same between towns in the median seat as it is between seats. Thus if 40% of towns in the median seat have 55% or more Labour voters, 40% of seats must have 55% or more Labour voters. Fractality therefore requires definition (a) and excludes definition (b.)

7 Let  $x$  be normally distributed with variance  $\text{var}(x)$ . Now let a set of  $y$ 's be normally distributed around each  $x$  with variance  $\text{var}(y)$ . It is a standard proof that the overall variance of all the  $y$ 's taken together will be  $\text{var}(x) + \text{var}(y)$ . Hence, if a set of  $z$ 's were normally distributed around each  $y$  with variance  $\text{var}(z)$ , the overall variance of all the  $z$ 's taken together would be  $\{\text{var}(x) + \text{var}(y)\} + \text{var}(z)$ ; and so on. If all the variances are the same (our fractal case) then  $\text{var}(x) + \text{var}(y) + \text{var}(z) + \dots + \text{var}(n) = n \cdot \text{var}(x)$ .

8 We could certainly think of an individual voter as 57.2% likely to vote Conservative. But this gets us nowhere until we make an assumption about how individual probabilities change when opinion moves nationally. If the 50%-likely Conservatives become 60%-likely, do the 30% Conservatives become 40%-likely, 36%-likely, or some other figure? Since we are looking at the consequences of a fractal electoral geography, we would have to make whatever assumptions keep the picture a fractal one. This would come to exactly the same thing as the approach we actually take.

9 When Labour has 50% of seats and votes,  $S_A = S_B = V_A = V_B = 0.5$ . Hence

$d(S_A/S_B)/dt = 4 dS_A/dt$  and  $d(V_A/V_B)/dt = 4 dV_A/dt$ . Hence

$$\frac{dS_A/dt}{dV_A/dt} = \frac{d(S_A/S_B)/dt}{d(V_A/V_B)/dt} = \frac{d(V_A/V_B)^n/dt}{d(V_A/V_B)/dt} = n(V_A/V_B)^{n-1} = n.$$

10 “Fractal n” as calculated this way has deviated by only tiny amounts from the values reported in table 1 within countries for elections since 1960.

11 first-preference votes

12 In 1996 New Zealand switched to a multi-member proportional representation system.

13 We have counted “seats” at 51 in the U.S.A. because we are looking at Electoral College votes for the Presidency. Since all the electors in any state vote the same way (bar two states), it is as though there are only 51 members of the Electoral College, but with several votes each (and some with more votes than others). Note that, while unequal electoral districts may well introduce bias into the system, they will not affect the value of n. To see this imagine that a uniform swing of x% across electoral districts causes y% of seats to change hands. Now imagine that one of the districts loses all but 10 of its voters, but that its proportionate political composition is the same as before. A uniform x% swing will still cause y% of seats to change hands.

14 Standard Errors are in parentheses.

15 The Durbin Watson test for serial correlation.

16 first-preference votes.

17 first-preference votes.

18 The 1993 observation was found to be an extreme outlier and was omitted from the sample.

19 The 1986 election in which the French ‘experimented’ with Proportional Representation was omitted.

20 In the 1996 election New Zealand switched to a MMP electoral system and so we exclude this observation.

21 UNR and Independents (until 1967), UDR and Independent Republicans (1968 & 1973), RPR and UDF (1978 onwards.) Since none of these pairs has literally been a coalition, but rather two distinct parties, which have fought against one another in general elections, we might expect them to get fewer seats per vote than a single right-wing party would. However, our figures refer to the second ballot in each French election, in which the two candidates with the most votes run off against each other. It has been rare for these two to be the two main parties of the right, and rarer still for one not to stand down in favour of the other. To all intents and purposes, therefore, their combined second-

ballot electoral performance has been equivalent to that of a single party, and they can be treated as such for the present purpose.

22 The choice of 100 is arbitrary, and a different base would give different values of k for given values of n.

23 The formula is  $n = \sqrt{\frac{1 - k^{\log_{100} V}}{1 - k^{\log_{100} S}}}$ , given our choice of 100 as the scaling ratio. Of course, a non-fractal electoral map is never going to be as neatly non-fractal as this. The formula above, though, does provide us with a rough index of over- or under-fractality when n, V and S are known.

24 P.Norris and I. Crewe, 'Did the British marginals vanish? Proportionality and exaggeration in the British electoral system revisited', *Electoral Studies*, 13 (1994), 201-221

25 This is confirmed by looking at the UK's Euro-elections of 1979, 1984, 1989 and 1994. Whereas "fractal n" remains steady, varying only between 1.949 (in 1989) and (1.933) in 1994, "actual n" is consistently above fractal n but on a downward, if irregular, trend, 2.951 in 1979, 3.043 in 1984, 2.103 in 1989 and 2.668 in 1994.

26 Calculated from figures in HMSO, *New Earnings Survey 1997* (London : HMSO, 1998).



27 D.Butler and D.Stokes, *Political Change in Britain*, (London: Macmillan, 2<sup>nd</sup> edn., 1974) chapter 6, provides a variety of evidence for this proposition, while William L. Miller, *Electoral Dynamics* (London: Macmillan, 1977) concludes that the social composition of your constituency affects your vote more than your own class does.