Pergamon

Political Geography 21 (2002) 1–31

Political Geography

www.politicalgeography.com

The 2000 Annual Political Geography Lecture

Manipulating maps and winning elections: measuring the impact of malapportionment and gerrymandering☆

Ron Johnston *

School of Geographical Sciences, University of Bristol, Bristol BS8 1SS, UK

Abstract

Geography is central to the operation of almost all electoral systems, through the interaction of two maps—the punctiform distribution of voters (and their political choices) and the territorial division of national space into constituencies. This interaction invariably results in election outcomes that are both disproportional and biased—with the allocation of seats being unequal to the distribution of votes across parties (and also across time for the same party). Such disproportionality and bias can be generated through the partisan strategies of malapportionment and gerrymandering, but in addition similar results can emerge when the procedure for defining constituencies is non-partisan. This paper argues that understanding the partisan and non-partisan nature of districting and its impact on election results requires formal

* Tel.: +44-117-928-9116; fax: +44-117-928-7878.
E-mail address: r.johnston@bristol.ac.uk (R. Johnston).

☆ Text of the Political Geography Specialty Group lecture delivered at the AAG Conference, New York, 1 March 2001 Much of my early work on elections was undertaken with and greatly stimulated by, Peter Taylor and it is a pleasure to be able to record my gratitude to him for his academic and personal companionship then and since. After his academic interests shifted to other subjects, most of my subsequent work has been undertaken with a range of collaborators, notably David Rossiter, Charles Pattie and Danny Dorling, but also Alan Hay, Ed Fieldhouse, Andrew Russell, Helena Tunstall, Iain MacAllister, Andrew Schuman and Dave Cutts. Over more than two decades we have worked on a variety of major themes, notably spatial variations in voting patterns at a variety of scales, the impact of local campaigning on election results, and the UK's redistricting procedures. The work reported here draws on all of that research, bringing it together in a way which illustrates the key roles of geography (or space) in elections. It would not have got anything like as far and as sophisticated as this without their assistance, which I very gratefully acknowledge here.

0962-6298/01/$ - see front matter
PII: S0962-6298(01)00070-1

measurement procedures and illustrates one such procedure, using the UK over the period 1950–1997 as its main exemplar.

Keywords: Bias; Elections; Gerrymandering; Malapportionment; UK

Introduction: on maps and measurement

Geography is central to the operation of virtually every electoral system used in representative democracies. Apart from the small number of countries—notably Israel and the Netherlands—which have a single constituency, divisions of the national territory are used for the key task of translating votes into seats.[1] In some, the goal is to achieve near-proportional representation, with each party's share of the legislative seats being equivalent to its share of the votes cast, subject to certain constraints. But in others—notably those linked to the Westminster system of government—different representation principles are employed; proportional representation is neither an explicit nor an implicit goal—and is very rarely achieved. Indeed, countries with that system of government using single-member constituencies to elect all of their MPs produce the majority of the most disproportional election results, with parties almost invariably having vote shares very different from their seat shares. The reason for this, it is widely appreciated, is the key role of geography in the process of translating votes into seats.

Despite that appreciation, however, relatively little research has been done by geographers on either how the manipulation of space in the translation process operates in different contexts or how extensive its impact is. There are significant exceptions to this, notably Graham Gudgin and Peter Taylor's (1979) seminal book on *Seats, Votes and the Spatial Organisation of Elections*, but very little subsequent work has built on their foundation. This lecture uses a case study of the UK's electoral system over the period 1950–1997 to take that work forward and show not only how deeply geography is implicated in the electoral process there, but also how appreciation of that fact by political parties has had a crucial impact on recent British election results, most notably the New Labour landslide of 1997.

The empirical study is fascinating in and of itself for the insights it throws on both electoral geography and the British electoral scene over half-a-century.[2] But there are more general themes to be drawn from it related to current trends within geography which have downgraded (and in some cases denigrated) the search for generalisation and order which characterised much research only a few decades ago (on which see Johnston, 2000). There are many aspects of our world—and elections is certainly one of them—whose study involves investigating the behaviour of very

[1] And in Israel, recent changes to the operation of the electoral system have allowed an important geographical element to be introduced by parties/candidates: see Hazan (1999).

[2] Only a small part of the results is reported here: further details can be found in two recent books—Rossiter, Johnston and Pattie (1999) and Johnston, Pattie, Dorling and Rossiter (2001).

large numbers of individuals treated not as individuals but as category members (in the examples used here, of people living in defined territories—Parliamentary constituencies in the UK case). One of our concerns should be to understand how that categorisation is produced, and what impact it has on the operation of representative democracy—in particular of who wins and loses elections, and why.

Maps, of two main types, are central to this. First there are the maps of people's political preferences—of the candidates and parties that they support. Secondly, there are the maps of the territories into which they are categorised—the constituencies which elect legislature members, from whom in many cases governments are formed. And the interaction of those maps is crucial: where the constituency boundaries are placed can have a telling effect on which parties win which seats; and once the boundaries are in place individuals and those who campaign for their votes may change their behaviour in order to influence election results.

Studying those maps and their interactions is not new, of course, and we have plenty of appreciation of particular cases. But that is all they are—particular cases. We know a lot about gerrymandering of various types in the USA over recent decades, for example—whether it be of how the different parties sought to influence the composition of a State Legislature (Morrill, 1973), for example, or how a single Congressional District was configured in Georgia to create a black minority–majority district (see, for example, Lennertz, 2000). But we lack any further appreciation of how widespread such practices are and what their impact on election results has been. Have such gerrymanders advantaged one party over another, and to what extent? This is the sort of question addressed here, in a case study of the UK over a 50-year period embracing 14 general elections. To what extent have election results been influenced by the long-term geographies of electoral support for the various political parties, interacting with the shorter-term geographies of constituency configurations within which that first geography is translated into Parliamentary representation? This calls for measurement, for understanding aggregate patterns within which individual behaviour is encapsulated.

Geography and the translation of votes into seats

The nub of the problem addressed here is summarised by a simple ratio derived from an election result: the seats:votes ratio—a party's percentage of the seats won in the election divided by its percentage of the votes cast. A ratio of 1.0 indicates proportional representation; anything else indicates that it is either under- or over-represented.

Fig. 1 shows seats:votes ratios for the UK's three main political parties at the 14 elections since 1950, the first when the current arrangements for defining constituencies were used. Three salient conclusions can be drawn. First, the Conservatives' ratio exceeded 1.0 at 12 of the 14 elections, and exceeded 1.2 (i.e. at least 20% over-representation) at a sequence of five between 1974 and 1992. Second, Labour's ratio was over 1.0 at most elections, including all those post-1959, but it rarely exceeded 1.2—although it achieved the highest for both parties (in 1997). Finally,

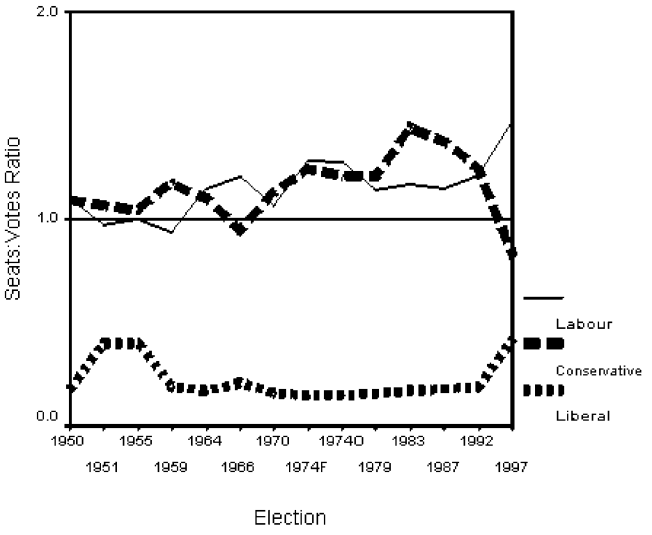

Fig. 1. Seats:votes ratios

the Liberals' ratios were always well below 1.0. Over-representation was the norm for the two larger parties, alongside under-representation for the smallest.

Why the apparently different treatment of the two largest parties, and in particular the major divergence between their ratios in 1997, when Labour achieved its highest ratio and the Conservatives their lowest? In 1979, the Conservatives won 43.9% of the votes cast and 53.4% of the seats. Four years later, their vote share was slightly smaller at 42.4% but their seats share increased to 61.1%. In 1987, their vote share increased slightly (to 43.4%) but their seats share fell (to 57.8%) and then in 1992 the vote share fell back slightly to 42.3% and the seats share to 51.6. Finally, after four Conservative victories Labour won in 1997 with 43.3% of the votes—and 63.6% of the seats. On five successive occasions, the winning party obtained between 42.3 and 43.9% of the votes—but between 51.6 and 63.6% of the seats.

From disproportionality to bias

These data provide not only clear evidence of disproportionality in UK election results, but also strong suggestions both that the parties are treated differentially and that the nature of this treatment changed over time.

Analysing seats:votes ratios cannot assist further attempts to sustain these inferences, because they do not compare like-with-like, either at any one election or across elections. For example, in 1997 Labour won 63.6% of the seats with 43.3% of the votes whereas 31.7% of the votes yielded 25% of the seats for the Conservatives. If the vote shares had been reversed, would the seats shares have been also—in which case the two parties would have been treated the same? If not, then the system is presumably biased against one of the parties, relative to the other. This is the basic definition of bias adopted in studies of the seats:votes relationship in the

US, as in early studies by Grofman (1983) and King and Browning (1987), following the pioneering work of Tufte (1973).

These questions were also addressed in a pair of pioneering papers by a New Zealand political scientist, Ralph Brookes (1959, 1960), which have unfortunately attracted very little attention. We have adapted Brookes' approach rather than those advanced by King, Grofman and others for a variety of reasons. The first is both technical and presentational: Brookes' method uses an easily-appreciated metric. The second, and more important, is that it both takes into account a range of factors that might contribute to the observed bias and can be decomposed (again, using the easily-interpreted metric) to show their relative impact. The US studies assume a two-party system and also implicitly assume that there is one dominant contributory factor in the production of bias—gerrymandering, either deliberate or non-partisan. They are able to demonstrate the impact of gerrymandering practices (as, for example, in King, 1989), but do not partial out the impact of malapportionment.[3] Grofman, Koetzle and Brunell (1997) have sought an integrated method of measuring bias which takes account of both malapportionment and gerrymandering (plus differences in turnout, a form of reactive malapportionment; see below); we have reanalysed their data using the method employed here (Johnston, Rossiter & Pattie, 1999).

Our applications of Brookes' method involve calculating what an election result would have been if the two leading parties had achieved the same votes share. It makes a key assumption, that change in a party's share of the votes between elections tends to be uniform across all constituencies, whatever its performance at the first of the two contests. This is a reasonable first approximation in the UK case, although there is the potential for experimentation with alternatives, such as the stochastic element introduced by King. Our analyses are thus based on reworking the results of the 14 elections so that the Conservative and Labour parties got an equal share of the votes at each—by reducing the leading party's share by a standard number of percentage points and increasing the other's share accordingly. Thus in 1997, Labour's share was reduced by 5.8 percentage points in every constituency whereas the Conservatives' share was increased by that amount—each then had 37.5% of the national total. That equality of vote shares was not matched by an equality in the allocation of seats—Labour would have won 82 more seats than the Conservatives (out of a total of 659). All of the analyses reported here discuss the bias at that one, hypothetical, situation—equal vote shares between the two main parties; they assess the difference between the two parties (Conservative and Labour) in the translation of votes into seats in that situation only.[4]

[3] In Gelman and King's (1994) important paper, for example, the trends in their bias measure (Fig. 1, p. 540) show bias substantially reduced at the beginning of each decade and then increasing steadily until the next round of reapportionment. This is entirely consistent with what one would expect from malapportionment: over time, population moves away from the cities where the Democratic party is strongest, meaning that increasingly it benefits from 'creeping malapportionment' whereby its seats become smaller whereas it's opponent's become larger.

[4] Grofman et al.'s (1997) method estimates the bias at the actual election result, as does Gudgin and Taylor's (1979) Gelman and King (1994) evaluate it over a range of situations, with each party's share of the two-party vote varying between 0.45 and 0.55. See also Jackman (1994).

Trends in the volume and direction of that bias across the 14 elections are shown in Fig. 2, in which a positive bias indicates an advantage to Labour and a negative figure a bias to the Conservatives. There was a major shift over the 50 years. The first five elections (1950–1964) produced a substantial pro-Conservative bias; there was generally very little bias at the next seven (1966–1987), save in February 1974 and 1979, when there was an advantage of some 20 seats favouring Labour; and finally the two 1990s elections produced very strong pro-Labour biases—which increased by some 40 seats over the previous contest on each occasion. The goal of the remainder of this discussion is to appreciate the origins of these changing biases.

The production of electoral bias

There is general appreciation of how disproportionality can be produced in single-member, first-past-the-post electoral systems dominated by two political parties, gained largely from our appreciation of two electoral cartographic abuses that have been widely practised in the USA—malapportionment and gerrymandering. In each, the goal is to maximise a party's returns from the votes that it obtains—to get as high a seats:votes ratio as possible.

Malapportionment

This produces biased results when (a) there are substantial variations in constituency size (measured as either population or electorate, depending on the local situation) and (b) one of the parties is strongest in the smaller constituencies (it is the beneficiary) whereas the other is strongest in the larger seats. This can come

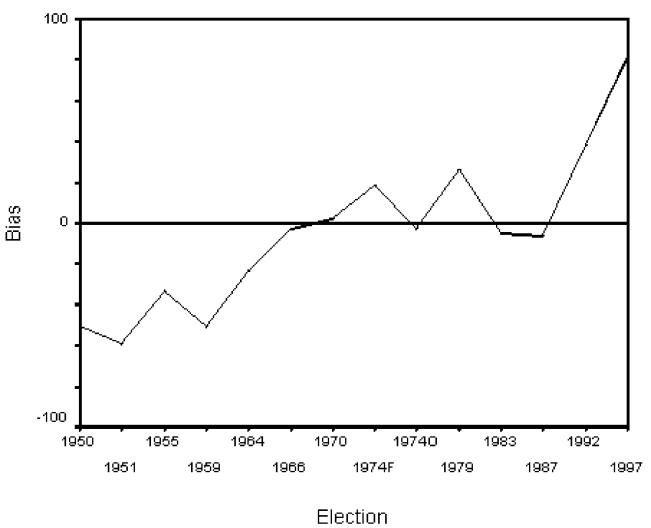

Fig. 2. Bias 1950–1997(with equal vote shares).

about either through *deliberate intent* if one party controls the districting process—creating larger constituencies in the areas where one's opponent is strong—or through *creeping malapportionment*, whereby changes in constituency size over time create smaller seats where one party is strong.[5]

Gerrymandering

This involves careful drawing of constituency boundaries by a party so that either it wins a particular seat or, more generally, it wins more seats than its opponent. There are two main types of gerrymander. In a *stacked (or packed) gerrymander*, the constituencies are arranged so that the opposing party wins a small number of constituencies by large majorities—and has little chance of winning any of the others, even if it benefits from a substantial shift of voter support. In a *cracked gerrymander*, on the other hand, the party controlling the boundary-drawing process creates as many seats as possible in which it has a majority—although this may be potentially damaging if the majorities are too small and the party loses votes to its opponent.

Both malapportionment and gerrymandering are used as strategies to promote a party's electoral advantage—Morrill's (1973) study of plans produced by the two parties in Washington State exemplifies such gerrymandering, for example. But such an advantage can emerge even when there is no deliberate intent, because although the districting is undertaken by independent bodies the geography of support for the two parties means that one is advantaged over the other. The existence of these non-partisan effects—particularly what they call non-partisan gerrymanders—was clearly demonstrated by Gudgin and Taylor (1979), and recent innovative simulation work by Cirincione, Darling and O'Rourke (2000) has provided a mechanism for identifying the likelihood of a particular result occurring by chance.

Apart from these two well-known 'abuses', there is a third set which is largely consequential upon the definition of constituencies (although it would be possible to incorporate them in the process given certain assumptions). We call these *reactive malapportionment*, because their impact is exactly the same as that of 'classic' malapportionment. Even with equal-sized constituencies, the equivalent of malapportionment can result if there is differential turnout and one party is strongest in the areas where abstention rates are greatest. The larger the number of *abstentions*, the smaller the number of votes needed to win a constituency; the winning party in areas of low turnout gets a better return from its votes because, in effect, it is strongest in an area of smaller constituencies. In exactly the same way, the larger the number of votes

[5] This was the situation in the classic cases charging malapportionment in the USA in the 1930s–1960s, when population decline in rural areas created small constituencies in areas of Republican strength whereas growth in urban areas resulted in much larger constituencies where the Democratic party was generally strongest. As a result, it took many more votes for the Democrats to win a seat than was the case for the Republicans, with the latter getting a much more favourable seats:votes ratio. The 'abuse' charged by plaintiffs was that an absence of redistricting denied them their equal rights enshrined in the Fourteenth Amendment—and it was eventually removed by the series of Supreme Court decisions that began with *Baker v Carr* in 1962.

for *third parties* (i.e. other than the two that dominate the system as a whole) the smaller the number of votes needed to win a seat, which will advantage one of the two main parties if it is strongest in the areas where the third parties are most successful also. Thus one of the two main parties will get a better seats:votes ratio than the other if its best performances are in either the constituencies where turnout tends to be lowest or third parties win most votes, or both—unless the third parties perform well enough to win some of the seats (which abstainers never do!).

Wasted, surplus and effective votes

These three types of electoral 'abuse'—all cartographic 'abuses', in that they are created by the act of drawing constituency boundaries in particular places—can be further appreciated by categorising each party's votes into three types: wasted, surplus and effective.

Wasted votes deliver no representation for a party, because they are obtained in constituencies that the party loses. *Surplus votes* similarly deliver no representation, because they are additional to requirements in constituencies that it wins. For example, in a constituency with 20,000 voters, A wins the support of 15,000 and B the remaining 5000. All of B's 5000 votes are wasted. Of A's 15,000, 5001 are need to defeat B and the remaining 9999 are surplus; the 5001 needed for victory form A's *effective votes*.

Where a party controls the districting process, its goal is to minimise its surplus and wasted votes—and thus maximise its effective vote proportion. Malapportionment, and reactive malapportionment, allow this: other things being equal, the smaller the average constituency that you win, and the larger the average constituency that your opponent wins, the smaller your number of effective votes per seat won. So does gerrymandering: a stacked gerrymander for your opponent means that it piles up large numbers of surplus votes, for example, and has a small effective vote percentage; a cracked gerrymander in your own favour means that you have few surplus votes while your opponent garners a large number of wasted ones.

Figs. 3–5 show the trends in surplus votes per seat won, wasted votes per seat lost and effective vote percentages at UK elections between 1950 and 1997, when the two main parties have equal vote shares. Except in 1970, Labour always amassed more surplus votes per seat won (Fig. 3), especially in 1983 and 1987—the two Thatcher landslide victories—which were Labour's poorest performances during the period, with almost all of its constituency victories in its northern, urban heartlands (Johnston, Pattie & Allsopp, 1988). The gap closed in the 1990s, but was still substantial in 1997.

The Conservatives have also generally been advantaged in wasted votes per seat lost, although there was virtually no difference between the two parties over the first five contests plus the two in 1974 (Fig. 4). But in 1997 the situation was substantially reversed; the Conservatives on average wasted 1500 votes more per seat lost than Labour.

The implication of these two trends, that a smaller percentage of Labour's than the Conservatives' votes were effective over the period, is clearly confirmed by Fig.

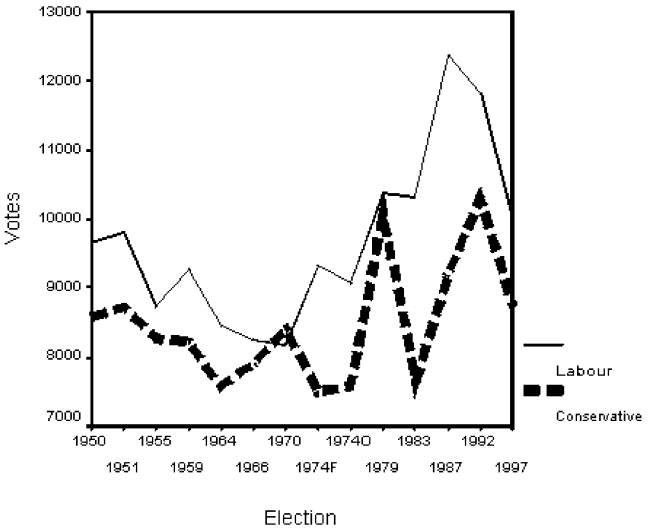

Fig. 3. Surplus votes per seat won (with equal vote shares).

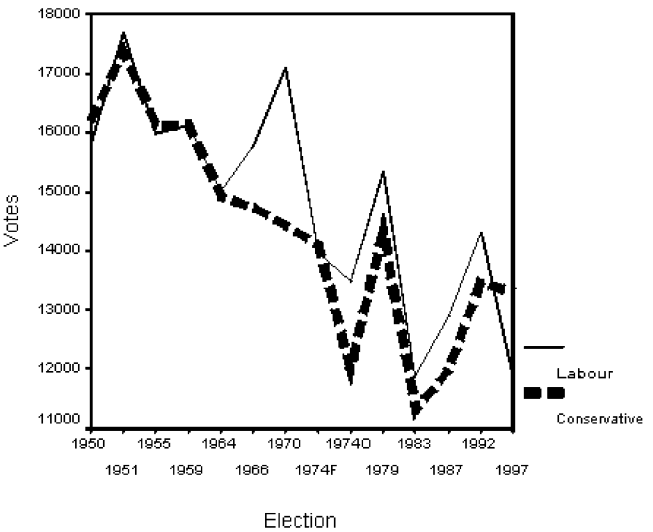

Fig. 4. Wasted votes per seat lost (with equal vote shares).

5—except in 1997. Up to 1979, the Conservative effective percentage averaged four points larger than Labour's (equivalent to about half-a-million votes). In 1983 and 1987 the gap was 9.7 points; it then narrowed to 4.7 points in 1992 before the turn-round in 1997, when Labour had 38.7 per cent of its estimated 11.5 million votes effective compared to the Conservatives' 37.3 with the same vote share.

For most of the period, therefore, the UK electoral system was biased against

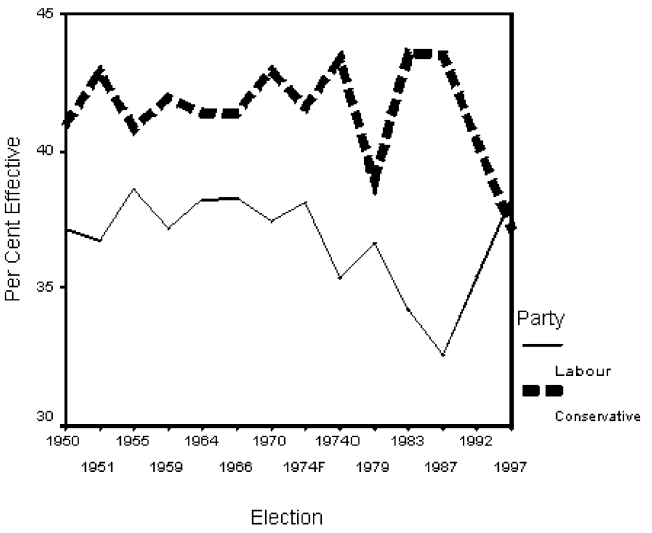

Fig. 5. Effective vote percentages (with equal vote shares).

Labour, whose votes were less effectively translated into seats than were their opponents'. This implies pro-Conservative malapportionment and/or gerrymandering. But the redistribution process is very largely non-political, so any cartographic 'abuses' must have been non-partisan outcomes of independent decision-making. Why and how? And why was 1997 so different?

Redistribution in the UK

The UK's map of Parliamentary constituencies has been redrawn on only nine occasions since the 1832 Great Reform Act.[6] The cartography was very much a political exercise on the first four—after each of the 19th century Reform Acts (1832, 1867, 1885) and in 1917—with the parties strongly influencing the outcome through close political involvement and direction. In 1944, however, Parliament instituted a system of regular reviews of the number and boundaries of constituencies, to be undertaken by independent Boundary Commissions (each chaired by a senior Judge although the Speaker of the House is the nominal chairperson) operating under rules laid down by Act of Parliament. The initial review was reported in 1947, with the recommended constituencies first used at the 1950 election; the subsequent reviews reported in 1954, 1969, 1983 and 1995.

The rules set out in the *House of Commons (Redistribution of Seats) Act 1944*, were amended on several occasions, restructured in the *Parliamentary Constituencies Act, 1986* and finally amended (on the timing of reviews only) in the *Boundary*

[6] The Boundary Commission for England announced the start of the tenth revision in February 2000.

Commissions Act, 1992; reviews are currently conducted every 8–12 years. They stipulate that each constituency should return one MP only, guarantee minimum numbers of seats for Scotland (71) and Wales (36),[7] and specify both a minimum and a maximum for Northern Ireland (16-18); no number for England is indicated, but Great Britain's total should not be 'substantially greater or less than' a given figure.[8]

The Commissions are required to respect organic and arithmetic principles when undertaking redistributions. Regarding the *organic*, they should allocate seats to the major local government areas (Counties, Unitary Authorities and London Boroughs), which are assumed to represent local communities, and when considering changes should take into account both the disruptions that these may cause and the local ties that may be broken.[9] The *arithmetic* principle is incorporated in the rules by a requirement that the Commissions create constituencies, within the local government template, whose electorates are 'as near the electoral quota as is practicable'. They have flexibility to depart from these rules if they deem it necessary, however—a flexibility considerably strengthened by a (to some dubious) Court judgement in 1982 (Rawlings, 1988). The basic purpose of the regular reviews is to counter any inequality in constituency electorates resulting from population movements, recommending that where necessary Parliament introduces new constituencies which meet the arithmetic principle within the constraints of the organic.

Although each Commission operates slightly differently, basic procedures are the same. The first step is defining an electoral quota by dividing the current national electorate by the number of seats. This is used to allocate seats to each local government area and determine whether it is necessary to combine adjacent areas to avoid inequalities. Each constituency is then defined as a group of contiguous wards (electoral areas for local governments). Provisional recommendations are published and representations invited—both positive and negative. If negative comments are received, a public Local Inquiry is convened, with an Assistant Commissioner hearing evidence, almost certainly including suggestions for alternative constituency configurations; the political parties dominate the proceedings in most cases (both by either supporting the Commissions' recommendations or promoting their own alternatives and by cross-examining other witnesses). The Assistant Commissioner recommends whether, and if so how, the Commission's recommendations might be altered—usually because of arguments made with regard to the organic rather than the arithmetic principle—and the Commission responds. (Commissions generally accept the Assistant Commissioners' advice.) The revised recommendations are sub-

[7] The *Scotland Act (1998)* significantly amended the *Parliamentary Constituencies Act (1986)* with regard to the number of seats for Scotland from the Fifth Periodic Review onwards. No minimum is guaranteed, and the Scottish Commission must use the same electoral quota to determine the number of seats in its Fifth Review as the English Commission.

[8] The figure was initially 591, and is now 613. For a full discussion of redistribution in the UK see Rossiter et al. (1999).

[9] There is no requirement to use the major local government boundaries in Northern Ireland and the requirement is weaker for Scotland than it is for England and Wales (and is very uncertain at present because of changes in Scotland's local government structure that have not been incorporated in amendments to the electoral legislation).

ject to further representations, but a second Local Inquiry is a rarity; the revised recommendations usually remain unchanged and become the final recommendations to Parliament.

Decomposing bias

Brookes' method has been adopted to explore how this independent, rule-bound redistribution procedure has generated the biases revealed in Fig. 2. His measure has the great advantages of not only of separating bias from disproportionality but also including a method of decomposing its easily-interpretable metric—the number of seats involved. Our adaptation of his algebra and formulae for the UK situation[10] identifies six separate bias components:[11]

Malapportionment	
1.	Inter-country variations in constituency size;
2.	Intra-country variations in constituency size;
Gerrymandering	
3.	The impact of the pattern and volume of surplus and wasted votes;
Reactive malapportionment	
4.	Abstentions;
5.	Third-party votes; and
6.	Third party victories.

Summing these six components irrespective of sign indicates the total amount of bias; taking account of sign gives a net figure. Total bias was approximately 50 seats at each of the first two elections, and then fell slightly before increasing steadily to over 120 by 1970 (Fig. 6). It fell to under 100 seats at the next three elections (still a substantial proportion of the total of 635 in 1979), followed by another run of increases to a total of nearly 180 in 1997. In net terms, bias at the first four elections was largely the same as the gross figure, before falling to almost zero for the seven elections between 1966 and 1987 (as indicated in Fig. 2) and then increasing substantially in the 1990s. The latter change was the result of an increased volume of pro-Labour bias (Fig. 7, which shows the sum of biases favouring each of the two parties). There were virtually no pro-Labour biases at the first four elections, hence the Conservative advantage then. From 1966 to 1987, the amount of pro-Conserva-

[10] In part building on sensible modifications suggested by Mortimore (1992).

[11] There is a seventh, relating to the particular situation in Northern Ireland, but it covers the elections up to 1970 only, after which none of the three main Great Britain parties contested seats there.

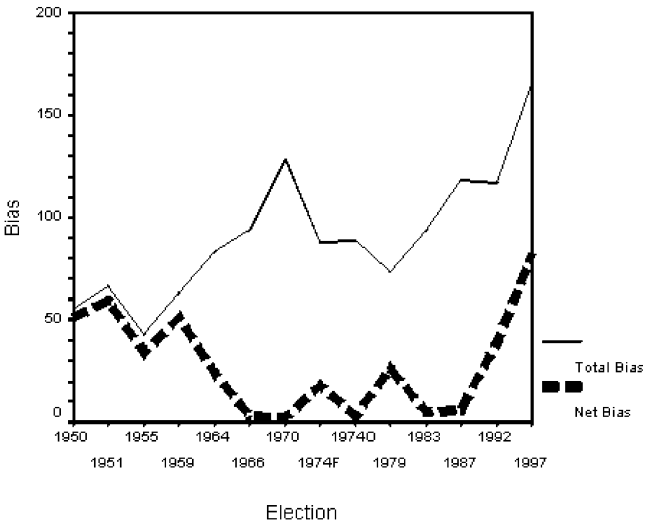

Fig. 6. Total and net bias (with equal vote shares).

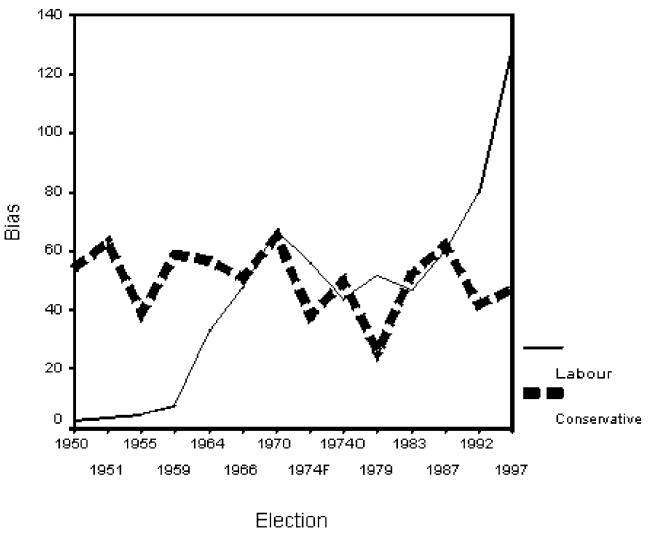

Fig. 7. Total bias to the two parties (with equal vote shares).

tive bias remained fairly constant but the pro-Labour bias increased, with the two virtually cancelling each other out. Finally, the last two elections produced a massive increase in the total volume of pro-Labour bias, from less than 60 seats in 1987 to over 80 in 1992 and then nearly 130 at the final contest in 1997.

Alongside changes in the size and direction of the bias were alterations in their

provenance (Fig. 8). The gerrymander component dominated the first few elections, accounting for just over half of the total bias (summed irrespective of sign). Its relative importance was more than halved by 1964, however, and it remained relatively insignificant throughout much of the rest of the period, though increasing again in 1997. Between 1950 and 1970 the two malapportionment and three reactive malapportionment components increased to provide about 40% of the total each. From 1970 on, the latter were generally the more important, accounting for over half of the total at the last two contests, although malapportionment was more important in 1979 and 1992 than at other elections—for reasons discussed below.

The production of non-partisan biases in the UK

An increasing volume of bias characterised the period 1950–1997, therefore and it increasingly favoured the Labour party as certain components declined in relative importance and others became more important; why? Redistribution in the UK is undertaken by independent, non-partisan Commissions operating a set of rules laid down by Parliament—although it is generally accepted that those rules are both ambiguous and contradictory and the 1982 Court judgement indicated that most were little more than guidelines only. The Commissions have considerable discretion in applying the organic and arithmetic principles, but there is no evidence that they do this in other than an entirely non-partisan way.[12] How, then, do they produce out-

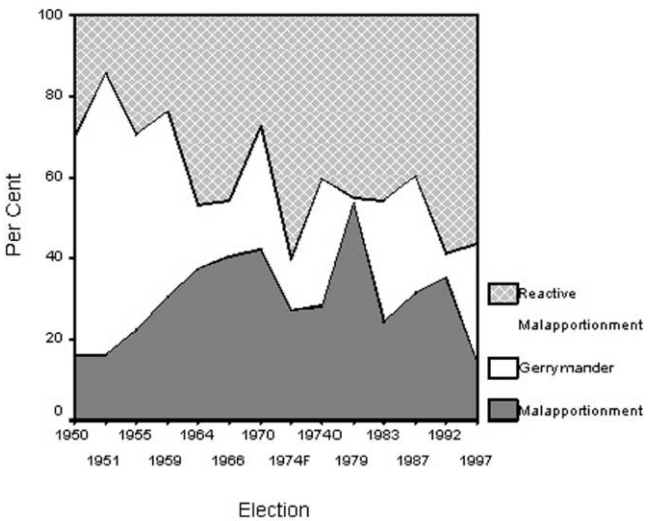

Fig. 8. Contributions of the three bias sources (with equal vote shares).

[12] We base this conclusion on interviews with the Commission members and many of the Assistant Commissioners during the period of the Third Periodic Reviews conducted by all four Commissions, 1991–1995. In England and Wales the political allegiances/activities of potential Assistant Commissioners are scrutinised closely—although errors occasionally occur.

comes that are so biased and why, if the system has remained largely unchanged over 50 years, has the volume and direction of those biases altered so considerably?

The basic answer to this question can be found in Graham Gudgin and Peter Taylor's (1979) pioneering work. They showed that disproportional treatment of one party over another would result from their geographies of support being overlaid by a constituency grid. Different geographies have different outcomes, depending on both the average support for a party across the constituencies and the skewness of its vote distribution. But different levels of disproportionality could be produced with the same geographies of support, depending on the detailed configuration of the constituencies employed. (Within cities, for example, constituencies organised zonally were likely to produce different results for the parties than constituencies organised sectorally.) The UK Boundary Commissions create constituencies as groups of contiguous wards[13] and a substantial number of configurations of wards meets the requirements of the rules for redistribution in each local government area. The Commissions—aided by their staffs of seconded civil servants—must choose from these combinatorial possibilities. Their selection process is non-partisan, but the outcome will almost certainly favour one party over another—one will get a better seats:votes ratio than the other at subsequent elections. The non-partisan procedure can, indeed almost certainly will, produce partisan outcomes.[14]

Disproportionality and bias result from the operation of geographical factors that are integral to the Boundary Commissions' method of working, therefore. But why has the volume and direction of the bias changed so markedly?

Malapportionment: UK style

The Boundary Commissions' regular reviews of constituencies are undertaken to reduce inequalities in their electorates. They never remove them entirely, however, and once new constituencies are created inequalities soon re-emerge.

Part of the reason why there is always some malapportionment element to the bias in UK election results is because two of the four countries are over-represented, and one of the two main parties—Labour—is now very much the stronger there. The pro-Labour inter-country malapportionment bias increased over the 50 years

[13] This is only a legal requirement in Northern Ireland but it has been the accepted, and largely unchallenged, modus operandi of all four Commissions throughout the period studied here. To a considerable extent, it is the only option open to them, since the electoral data on which they must base their decisions are not available for smaller areas. In some parts of each country, at some dates, there are two sets of electoral wards where local government is divided into two tiers. The Commissions use the smaller wards whenever possible.

[14] The Commissions repeatedly stress their non-partisan approach. In the volume produced to describe its work at the outset of the Fifth Periodic Review, the Boundary Commission for England (2000, 2) stated that:

> The Commission are an independent, non-political and totally impartial body. They emphasise very strongly that the results of previous elections do not and should not enter their considerations when they are deciding their recommendations. Nor do the Commission consider the effects of their recommendations on future voting patterns.

(Fig. 9). The 1944 decision to guarantee no less than 35 seats for Wales and 71 for Scotland was entirely political, taken to placate opinion there: despite some claims to the contrary, Scotland's guarantee was not part of the Act of Union in 1707 (see McLean, 1995, and Rossiter, Johnston & Pattie, 1997a).

The pro-Labour bias has increased for three reasons. First, because of the guarantee, the ratio of MPs to electors for Scotland and Wales has fallen relative to that for England, where population growth has been much greater. At the Initial Review in 1947, for example, the average Scottish constituency electorate was 49,620 (85% of the English average) and that for Wales 51,641 (88%); 50 years later, the respective percentages after the Fourth Review were 80 and 81. Scotland and, especially, Wales have become more over-represented, in part because of the second reason—each has since been allocated more seats: Wales now has 40 and Scotland 72.[15] Thirdly, Labour's electoral support has increased relative to the Conservatives' in the two countries. It always had a substantial lead in Wales and the difference between the two increasingly diverged in Scotland after 1959 (Dyer, 2001)—Labour had twice the Conservative vote share in both countries in 1997, when the Conservatives failed to win a seat in either. Compared to England, therefore, where apart from the elections in 1979–1992 the two parties vote shares were approximately equal, Labour has been substantially advantaged in Scotland and Wales by a combination of its electoral performance and the countries' over-representation. Malappor-

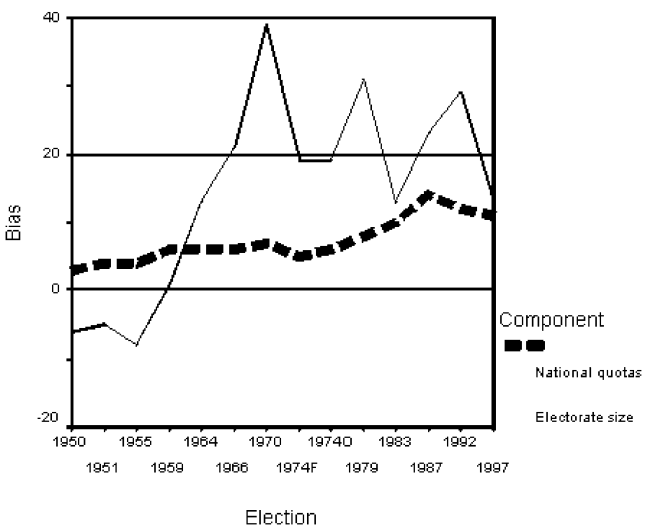

Fig. 9. The malapportionment bias components (with equal vote shares).

[15] These extra allocations result from the procedures operated by the Commissions (on which see Rossiter, et al., 1999 Johnston, et al., 2001). The increases were not absolutely necessary—the Scottish Commission determined that it would not increase the number during its Fourth Review—but the Welsh Commission has argued that it is very difficult to avoid them.

tionment was introduced for non-party political reasons, but the consequence favoured Labour.[16]

The second, and larger, source of pro-Labour malapportionment bias is differences in constituency electorates within countries (Fig. 9). This component's size and direction have varied considerably. Until 1959, the bias was pro-Conservative, because the English Commission (responsible for 80% of the constituencies) used the flexibility under the special geographical considerations rule to create smaller constituencies in rural areas on the grounds that these were more difficult for parties and MPs to service because of their areal extent and low population densities. From then on, however, the flexibility was applied very sparingly in remote rural areas only—only seven additional constituencies were created for this reason in the Third Periodic Review and two in the Fourth. Electoral equality was the dominant criterion from 1954 on, but a malapportionment effect emerged because population changes favoured Labour. The constituencies experiencing population decline between reviews were predominantly urban, and returned Labour MPs; the constituencies that grew were dominantly suburban or rural, with Conservative MPs. Thus the older the constituencies the greater the urban:rural difference in average constituency size, and the greater the pro-Labour bias.

The sawtooth trend in this pro-Labour bias from 1959 on reflects the impact of the reviews. Every time new constituencies were introduced—for the general elections in February 1974, 1983 and 1997—electorate inequalities were reduced and the pro-Labour bias fell. But the sawtooth effect is not regular: the pro-Labour bias was higher in 1970 (the last election before the new constituencies were introduced in 1974) than 1979, for example; this was for partisan reasons, and reflects political 'interference' in the work of the Boundary Commissions.

The Boundary Commission for England was about to report in spring 1969, and knowledge of its proposals through the public consultation process led Labour to calculate that it could lose some 20 seats at the next general election (due by 1971).[17] Labour was in power, and decided not to implement the recommendations. (The Commissions only recommend to Parliament, which has to debate but not accept them; it cannot amend them, but the Secretary of State can, before presenting them to Parliament.) Labour's leaders argued that forthcoming wide-ranging local government reforms would require a further review almost immediately. But the opposition insisted that the recommendations be placed before Parliament, and Labour's proposal that only the recommendations for Greater London (where local government reform was completed before the Commission sat) be implemented was defeated in the House of Lords. So the government introduced the full set of recommendations, but whipped its MPs into defeating them. The 1970 election was then held in the

[16] Labour's advantage will be cut in Scotland after the Fifth Review, because its Commission was instructed to use the same electoral quota as England—to reflect the significant devolution of power to the Scottish Parliament, first elected in 1999. But the constituencies created by the Review may not be sued in a general election until 2009, and there is no similar provision to remove Welsh over-representation.

[17] This was known because all of the recommendations had been subject to prior public consultation.

old constituencies, yielding a 39-seat malapportionment bias to Labour (which nevertheless lost the election).

The next set of Boundary Commission reviews was completed in 1982. Again, Labour expected to lose seats as a consequence, and wanted to delay their implementation. But it was in opposition and so before the English recommendations were delivered to the Secretary of State it filed a case claiming that the Commission had 'misdirected' itself and not fulfilled the requirements of a review; it could have produced greater electorate equality in a number of parts of the country. The claim was rejected by the Courts, and the 1983 election fought in the new constituencies; Labour had a malapportionment bias of 13 seats, instead of about 40 which may have been the case had the old constituencies (defined using 1965 electoral data) been used.

In 1992 the Conservatives won an unexpected fourth general election in a row. The Boundary Commissions were already conducting their Fourth Periodic Reviews, but did not need to report until at least 1995; the last possible date was 1998. The Conservatives realised that their chances of victory at the next general election would be increased if it were fought in new constituencies, and their *Boundary Commissions Act 1992* required the Commissions to report by December 1994. The new constituencies were in place for the 1997 election—which the Conservatives lost, but which had a pro-Labour malapportionment bias of 13 seats (compared to 29 in 1992).

Creeping malapportionment produces bias in the UK because Labour is electorally strongest in areas where population declines between redistributions. Both political parties have sought to influence the size of this bias component, therefore, by changing the timing of reviews—either directly (through legislation) or indirectly (by delaying their implementation). They have succeeded on two occasions, with clear consequences on the level of pro-Labour bias at subsequent elections.

Gerrymandering the UK

The gerrymander bias component favoured the Conservatives at 12 of the 14 elections—February 1974 and 1997 were the exceptions—although it was very small in 1979 and 1992 (Fig. 10). This has occurred because Labour's vote distribution is spatially more concentrated than the Conservatives', both across the country and within the individual local government areas to which constituencies are allocated. As a result, Labour tends to win by much larger majorities in its areas of strength than do the Conservatives in their heartlands, producing an infrastructural gerrymander whereby: (a) in the areas where Labour is the largest party, Labour tends to amass large numbers of surplus votes per seat won in the equivalent of stacked gerrymanders, with the Conservatives getting relatively few wasted votes per seat lost; and (b) in the areas where the Conservatives are the largest party, they amass many fewer surplus votes per seat won in what are the equivalents of cracked gerrymanders, whereas Labour gets a large number of wasted votes per seat lost.

Table 1 illustrates this very significant inter-party difference in the operation of the infrastructural gerrymander with the estimated results for two areas in 1983 when the two parties have equal vote shares nationally. Northamptonshire is an area of

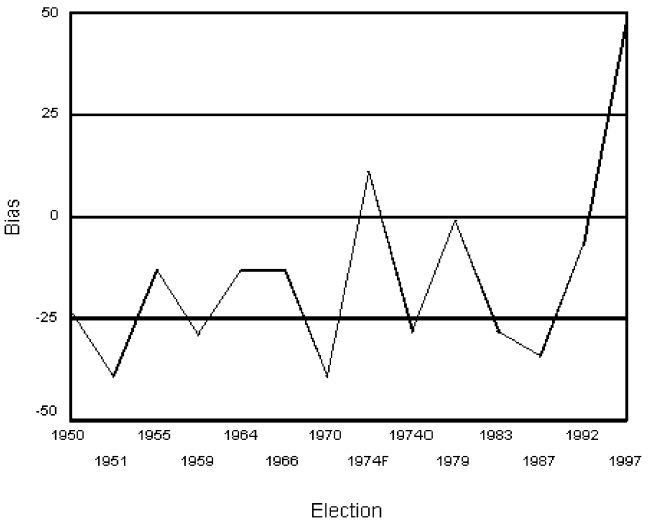

Fig. 10. The gerrymander bias component (with equal vote shares).

Table 1
Wasted votes per seats lost and surplus votes per seats won at the 1983 general election (when the Conservative and Labour parties have equal vote shares nationally): Northamptonshire and Sheffield[a]

	CVA	LVA	AV	CS	LS	CW	LW
Northamptonshire							
Corby	17,117	21,368	9905	0	4251	17,117	0
Daventry	22,568	13,628	13,221	8939	0	0	13,628
Kettering	19,544	13,797	14,637	4906	0	0	13,797
Northampton North	19,355	17,042	12,829	2312	0	0	17,042
Northanpton South	22,986	15,370	11,698	7615	0	0	15,370
Wellingborough	21,700	17,673	12,994	4026	0	0	17,673
Average				5560			15,502
Sheffield							
Attercliffe	8023	26,498	10,241	0	16,256	8023	0
Brightside	4535	28,884	10,322	0	18,561	4535	0
Central	4793	27,874	7969	0	19,904	4793	0
Hallam	22,834	14,479	15,077	7757	0	0	14,479
Heeley	11,706	28,148	12,813	0	15,314	11,706	0
Hillsborough	11,577	25,204	19355	0	5848	11,577	0
Average					15,185	8127	

[a] Key to columns: CVA, Conservative vote total (adjusted to ensure equal vote shares nationally); LVA, Labour vote total (adjusted to ensure equal vote shares nationally); AV, Alliance vote; Conservative surplus; LS, Labour surplus; CW, Conservative wasted; LW, Labour wasted.

Conservative strength; the party has 55% of the two-party vote and wins five of the six seats with majorities below 9000. In Sheffield on the other hand, Labour wins five of the six seats with 70% of the two-party vote; it averages 15,185 surplus votes per seat won (compared to the Conservatives' 5560 in Northamptonshire) and the Conservatives waste 8127 per seat lost (compared to Labour's 15,502 in Northamptonshire). The Conservatives get a much better return for 55% of the votes in their heartland than does Labour for its 70% (the seats:votes ratios are: Conservative in Northamptonshire, 1.52; Labour in Sheffield, 1.19).

This general pattern is repeated through the country, with Labour amassing surplus votes where it wins and wasted votes where it loses to a greater extent than the Conservatives. Thus the Conservatives get much higher effective vote percentages in areas where they are the strongest party than Labour does where it is. The geographies of both party support and Boundary Commission seat allocation interact to produce a pro-Conservative infrastructural gerrymander—except in February 1974 and, especially, 1997. What was different in the latter case?[18] How did Labour benefit from such a massive turn-round, from a bias component worth 34 seats to the Conservatives in 1987 to a pro-Labour advantage of 48 just 10 years and two elections later—albeit with a redistribution in between?

The first part of a four-part answer is that the *cracked gerrymander* worked to Labour's advantage in many of the areas of Conservative electoral hegemony prior to 1997. Conservative support fell by 9 percentage points between the 1992 and 1997 elections and the party lost many of the seats won at previous elections by comfortable but not large majorities. (Labour won five of the six in Northamptonshire, for example; it would have retained two with equal vote shares.)

The second reason is successful *gerrymandering-by-persuasion* by the Labour party during the Boundary Commissions' Fourth Periodic Reviews. Each local government area has a large number of possible configurations of contiguous wards into the required number of constituencies. Not all will produce the same electoral outcome: in Sheffield, for example, although most configurations would result in Labour winning five of the six seats, other outcomes were possible in a relatively small number of cases.[19] Thus if a party can convince the Boundary Commission, through the Assistant Commissioner, to adopt a configuration favourable to it electorally, it would achieve a local gerrymander. Labour realised this during the Third Periodic Review and developed a strategy to try and maximise its returns from the Fourth—for which their opponents were totally unprepared. A senior party staffmember identified a pro-Labour configuration for every area, and was helped by

[18] The focus here is on 1997 because the pro-Labour gerrymander bias was so large then, and because of the special circumstances of the February 1974 election, called by Edward Heath on a 'who rules?' platform after a sequence of miners' strikes which had drastic consequences for the British economy and society and when 'third parties' increased their vote share very substantially.

[19] This conclusion is based on research (following that of Gudgin & Taylor, 1979) using a computer program specially written to identify all possible solutions to the Boundary Commissions' problem. (For a summary, see Johnston & Rossiter, 1982. The later work of Cirincione et al., 2000, takes this approach forward.)

senior politicians in discussions with local party officials and members to ensure that they promoted it in representations to the Commission and in oral evidence at the Inquiry. Of course, they could not make the case on electoral grounds, and so employed whatever elements of the rules—usually those associated with the organic rather than the arithmetic principle—were most likely to sustain a convincing case. This strategy was very successful and it was widely recognised that Labour won 'the battle of the boundaries'. Nevertheless, the overall gains were fairly small; the party claimed that a net gain of ten seats (i.e. those it should win if they had been used in 1992), but our estimates suggest that this is rather large.[20] Labour's gerrymandering-by-persuasion worked, but not to the extent needed to achieve the large gerrymander bias recorded for 1997.[21]

The third reason is the impact of *tactical voting* by opposition party supporters in Conservative-held constituencies. Tactical voting occurs when supporters of party (A) decide to vote for their second choice (B) because its candidate stands a better chance of defeating the candidate of a third party (C): defeat of C has the highest priority for them, even if it means victory for B in that constituency rather than A.[22]

Increased support for the Liberal party (now the Liberal Democrats) after 1970 (and, to a lesser extent, for the nationalist parties in Scotland and Wales) meant that many constituencies either became three-way contests (with all three candidates having some real chance of success) or one of the leading two leading parties nationally was relegated to a poor third place, with the Liberal Democrats replacing it as the main opposition locally. These changes were concentrated in areas of Conservative strength, where the Liberal Democrats replaced Labour as the main challenger in many seats (especially in the south of England; Johnston et al., 1988); a Liberal Democrat was much less likely to come second to a Labour victor (especially in the north). Thus, in each election after 1979, when a main goal of the opposition parties was removing the Conservative government, tactical voting was a sensible strategy—especially in 1997 when the Liberal Democrats more firmly identified themselves as a left-of-centre party than previously.[23]

Successful tactical voting helps the major party involved increase the effectiveness of its votes. In a Conservative-held seat where Labour is in second place, for example, if voters who would otherwise have voted Liberal Democrat transfer that support to Labour this may help it to defeat the incumbent—probably by a small majority, thus generating relatively few surplus votes (but a large number of wasted votes for the Conservatives). Furthermore, in a Conservative-held seat where the

[20] See Rossiter et al. (1999), Chap. 8.

[21] One clear example of this gerrymandering-by-persuasion strategy is provided by the Fourth Review in Northern Ireland, where the 'pro-Republican' parties successfully engineered a, for them favourable, change to the Commission's provisional recommendations with regard to the number of seats having Roman Catholic majorities (Rossiter, Johnston & Pattie, 1998).

[22] The American term for this is strategic voting.

[23] There were no formal inter-party agreements on this, either nationally or in individual constituencies, although in a number of seats tactical voting was covertly encouraged. The general closeness of the two parties was illustrated a few months before the election, however, when a Joint Committee produced an agreed report on constitutional change.

Liberal Democrats are in second place and thus have the best chance of unseating the government's candidate, Labour voters transferring their support to the Liberal Democrats can have two effects. First, they may help the latter to win the seat—again probably with a relatively small majority, yielding large numbers of wasted votes for the defeated Conservatives. Secondly, because the Labour vote falls its number of wasted votes is reduced. Thus tactical voting aids Labour, not only in helping it to overhaul the Conservative incumbent in more seats than it might do without tactical support from otherwise Liberal Democrat voters, but also in increasing the effectiveness of its vote total—it wastes fewer votes where it loses and amasses only small numbers of surplus votes where it wins unexpectedly.

The extent of tactical voting can be established from analysing constituency flow-of-the-vote matrices. Data from national sample surveys show the volume and direction of vote-shifting between 1992 and 1997 (Table 2); among Conservative voters in 1992, for example, only 57.1% supported the party again in 1997, with a further 9.2% shifting to Labour and 5.4% to the Liberal Democrats. In Conservative-held seats differences in several of these cells should vary according to which party was best-placed to unseat the incumbent were tactical voting to occur. Where Labour was in second place, for example, the percentage of Labour loyalists should be higher than where the Liberal Democrats were lying second, along with the flows to Labour from those who voted either Conservative or Liberal Democrat in 1992 or didn't vote at all then.[24] Liberal Democrat loyalty, on the other hand, and the three sets of flows to it, should be higher in constituencies where that party's candidate was

Table 2
Inter-election flow-of-the vote, 1992–1997[ab]

	1997				
1992	C	L	D	N	NV
C	57.1	*9.2*[c]	*5.4*	0.3	25.9
L	1.6	*80.6*	*2.8*	0.5	12.5
D	3.2	*14.8*	*58.7*	0.6	19.8
N	30.	13.2	1.6	67.9	11.3
NV	11.5	*19.1*	*6.2*	1.1	57.9

[a] Key to rows/columns: C, Conservative; L, Labour; D, Liberal Democrat in 1992 and 1997); N, nationalist parties; NV, did not vote.

[b] The data are percentages of the row totals (i.e. those who voted for the relevant party at the first election of the pair).

[c] The figures in italic show the flows employed in the studies of tactical voting.

[24] The non-voting category includes both those who abstained in 1992 and those who were not registered as electors then; the two are not readily disentangled at the constituency scale.

second in 1992.[25] Finally, the flows between the two parties should reflect the local tactical situation; Labour-to-Liberal Democrat flows should be smaller where Labour is second, for example, whereas Liberal Democrat-to-Labour flows should be larger.

To test these expectations, entropy-maximisation was used to estimate the 1992–1997 flow-of-the-vote matrix for each constituency. (The method has been widely used for this purpose over the last two decades; see Johnston & Pattie, 2000.) Fig. 11 shows the average flow in each of the eight relevant matrix cells for Conservative-held seats according to which party was in second place. In every case, the difference was as predicted by the tactical voting hypothesis.

The potential for tactical voting success is greater in some constituencies than others, however; the larger the Conservative majority, the smaller the chances of the party's incumbent being unseated by the combined voting power of Labour and the Liberal Democrats. Thus, the more marginal the seat, the greater the potential from tactical success and the greater the probability of the relevant flows deviating substantially from the overall average. The more marginal the Conservative-held seat

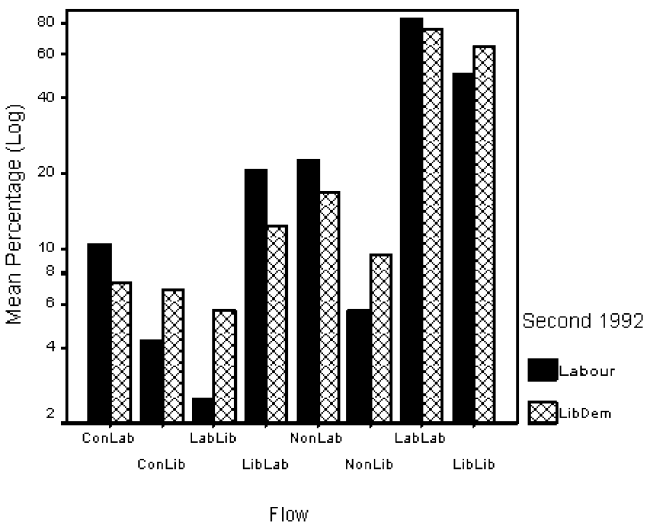

Fig. 11. Average tactical voting flows 1992-1997 in Conservative held seats.

[25] Because of the redistribution between 1992 and 1997, the latter election was fought in new constituencies. We have estimated what the 1992 result would have been if it had been contested in those constituencies, using ecological regression, in order to conduct these analyses (on which see Rossiter, Johnston & Pattie, 1997b). In addition, because of the large upsurge in Labour support nationally, there were constituencies where it notionally occupied third place only slightly behind the Liberal Democrats where it was perceived that Labour had the best chance of unseating the Conservative. (The Liberal Democrat vote share fell between 1992 and 1997, although its number of MPs doubled.)

the greater the volume of tactical voting, as shown in Fig. 12 for flows from the Liberal Democrats to Labour. The gap between Labour and Liberal Democrat seats in the size of this flow was greatest in the most marginal seats (i.e. where the gap between the first- and second-placed parties was close to zero percentage points): the wider the margin, and the larger the task of overhauling the incumbent, the smaller the gap (i.e. the less tactical voting that occurred), with the two trends converging on the safest Conservative seats.

This suggests a highly sophisticated electorate, with a substantial proportion of its members aware of, and acting upon, the tactical situation. That is inconsistent with general appreciations of voting behaviour, which suggests that collecting information with which to make such an informed choice is uneconomic for most electors. Instead, they are more likely to respond to stimuli, voting tactically when canvassed to do so. Tactical voting should occur in considerable volume where the parties campaign for it. The intensity of a party's constituency campaign can be indexed by its expenditure there, so that the more one party spends relative to the other, the better its relative performance should be. Fig. 13 shows this; the more that Labour spent relative to the Liberal Democrats in Conservative-held seats the larger the LibDem to Labour flow relative to the flow in the opposite direction. Campaigns provided information, increased the volume of tactical voting, and helped Labour improve the effectiveness of its votes in 1997.

The final part of the answer follows on from the last, emphasising the role of *targeted campaigning*. Conventional wisdom for some decades in British psephology has held that local campaigns have no influence—a conclusion based on little more than slight anecdotal evidence and increasingly challenged by research findings (e.g. Denver & Hands, 1997; Pattie, Johnston & Fieldhouse, 1995; Whiteley & Seyd, 1994). This work has shown that the Labour party has increasingly targeted its con-

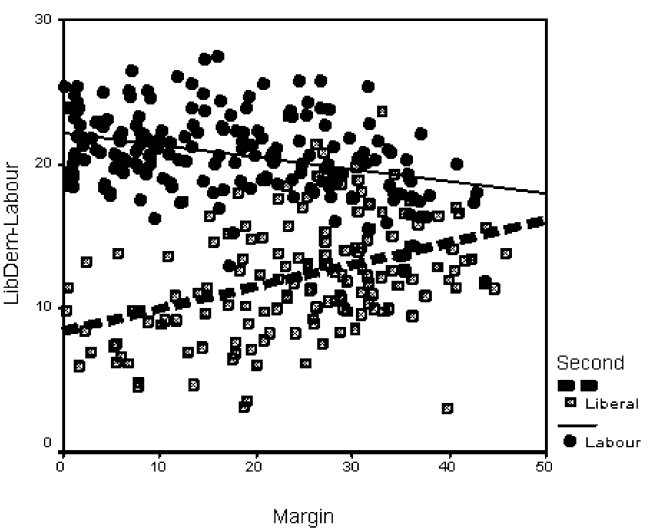

Fig. 12. Liberal Democrat to Labour flows and marginality, 1992–1997.

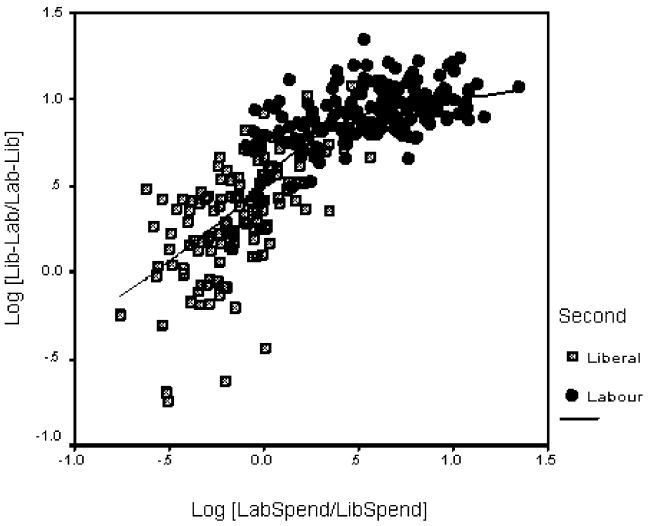

Fig. 13. Tactical voting and spending.

stituency campaigns on the seats that matter—the marginals that it might either win or lose with a relatively small shift in voter support. Thus it has paid little attention to either its safe seats, where additional surplus votes will bring no return in terms of seats won (and where abstentions have increased at recent elections), or those it has no chance of winning—where the effort of seeking additional wasted votes is not worth it. The Conservative party has been much less focused in this—in part because of the independence of its constituency parties, which makes the allocation of resources to desirable target seats difficult. Local campaigning brings dividends in terms of votes, which can be translated into seats if they are won in the right places—something that Labour (and the Liberal Democrats) have been much more adept at in recent years than the Conservatives.

Reactive malapportionment

Fig. 14 shows the trends in these three bias components. The two for minor-party votes and victories moved in opposite directions. As the number of minor party votes increased, so the Conservatives benefited more. Third-parties (especially the Liberal Democrats) performed much better in Conservative- than Labour-held seats, reducing the vote threshold for victory further in the former than the latter ands generating a pro-Conservative bias of over 30 seats in February 1974, 1992 and 1996. But as the Liberal Democrats won more seats, the Conservatives were disadvantaged, since most of the victories were in seats formerly held by the Tories. The pro-Labour third-party victories bias component was substantially less than the Conservatives gained from third-party votes, until 1997 when the number of Liberal Democrat victories doubled (even though the party's vote share fell!)—in part because of tacti-

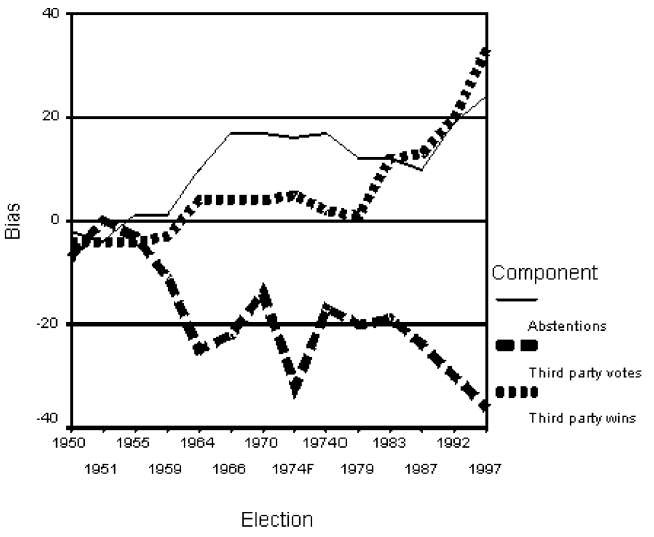

Fig. 14. Bias due to reactive malapportionment (with equal vote shares).

cal voting. Thus the net advantage to the Conservatives from the two components fell to just three seats—from a maximum of 27 in February 1974.

Finally, the abstentions component increasingly favoured Labour after 1955, being worth 24 seats in 1997. Turnout at British general elections fell from just over 80% in the early 1950s to a little over 70% by 1997. It has always been lowest in Labour-held seats, notably those in the inner cities where alienation from the political system tends to be greatest. During the 1980s and 1990s, as Labour increasingly focused its campaigns on marginal seats—both those it needed to hold and those it hoped to win—so less attention was paid to many of its inner city safe seats, as indicated by the low intensity of its campaigns there (why campaign hard for more surplus votes?). This has been to its considerable advantage, as Fig. 14 shows. Lower turnout in Labour's heartlands means fewer surplus votes per seat won and, with no incentive to increase it (there are no seats at stake, especially when, as in 1997, an overall Labour victory looks certain), the party has been happy to benefit from this.

Widening the applications

As stressed at the outset, this paper has focused on a case study of the 14 UK general elections between 1950 and 1997. But the method used is widely applicable, as can be briefly illustrated by the US Presidential election in 2000, when George Bush won 47.998% of the votes cast but defeated Al Gore (with 48.343%) by 271 to 267 votes in the Electoral College.[26]

[26] The number of votes for each State in the Electoral College is the same as its number of Representatives and Senators in the US Congress: all have two senators, but the number of Representatives is proportional to State population. Application of Brookes' method in this context involves dividing the

If the two candidates had tied, with 48.171% of the votes each, according to the measure of bias developed here, Bush would have won by 294 to 244 votes in the Electoral College: the shift of just 180,000 votes (out of over 103,000,000 cast) would have led to him winning three more states—Iowa, New Mexico and Wisconsin. As Table 3 shows, this pro-Bush bias of 50 seats in the Electoral College (out of 538) results largely from malapportionment (Bush tended to win in the smaller States with fewer voting-age residents per Electoral College vote) and a stacked-like gerrymander (Gore amassed more surplus votes per Electoral College vote won, largely through his substantial leads in the popular vote in both California and New York)—with the consequence that overall 47.5% of Bush's votes were effective, compared to 40.4% of Gore's. Only the pattern of third-party votes favoured Gore, to a small extent:[27] the geography of the election very much benefited his opponent, who won despite being the less popular candidate nationally. (For more details on this analysis, see Johnston, Pattie & Rossiter, 2001.)

Conclusions

New Labour's general election victory in 1997 in the UK was widely hailed not only as a landslide, but also as the outcome of a very disproportional translation of votes into seats. What was not widely appreciated, however, was just how much this landslide was a function of geographies, specifically the key interaction of two sets

Table 3
Bias in the result of the USA 2000 Presidential election

	Bush	Gore
The Bias Components		
Malapportionment	14.8	0.0
Gerrymander	38.5	0.0
Turnout	1.6	0.0
Minor Party Votes	0.0	2.2
The Bias Origins(with equal vote shares)		
Average voting age population per electoral college vote won	372,000	393,900
Average turnout (%) per electoral college vote won	52.2	54.3
Average minor party vote (%) per electoral college vote won	3.7	5.2
Average surplus votes per electoral college vote won	20,600	24,800
Average wasted votes per electoral college vote lost	82,740	80,676
Percentage votes effective	47.5	40.4

number of votes in each State for each candidate by the State's number of Electoral College votes, thereby increasing the number of 'constituencies' from 51 (the number of States, plus DC) to 538—the number of Electoral College votes.

[27] Even so, if—as seems very likely—most of the voters who favoured the most popular of the third party candidates (Ralph Nader, of the Green Party) had instead voted for Gore, he would almost certainly have won Florida and New Hampshire, and thus the Electoral College.

of maps—of party support and of constituencies—let alone how much that interaction resulted from political parties (mainly Labour) manipulating within the system to its advantage. The studies summarised here have provided that appreciation, and put the 1997 result into context—albeit using a set of measures based on a hypothetical situation (derived from the actual result) when the two parties have equal vote shares.

Providing that appreciation has involved measurement, the application of mathematical and statistical reasoning to understanding the nature and extent of bias. The procedure adapted to this purpose provides substantial insights—and there are clear opportunities both for developing it further and for exploring alternative methodologies—as briefly illustrated by the application to the US Presidential election in 2000. Several other procedures have been involved—including the entropy-maximising method that we have widely used over the last two decades to estimate unknown quantities such as spatial variations in flow-of-the-vote matrices, but which sadly have not been subject to detailed critical assessment let alone wider application.[28] Without such methods, will we ever be able to answer Rush's (2000) question—"would we recognise a gerrymander if we saw one?".[29]

Electoral and political geographers, like their contemporaries in most other branches of their discipline (and in other social sciences too), have been turning their backs on such technical issues in recent decades, with the consequence that we know an increasing amount about the parts, but relatively little about the wholes. Graham Gudgin and Peter Taylor showed us the way in the late 1970s, but very few have followed their signposts. This is disappointing, not only because of the inherent interest of the questions posed here regarding the nature and extent of bias but also because of the potential importance of the answers. The surge in the number of countries that operate a form of representative democracy following the 'end of history' in recent years has seen widespread experimentation with various electoral systems—alongside further experiments in other countries where disillusion with the results of one system has led to others being adopted in the hope that 'better politics' will emerge. Such experiments should be guided by knowledge of the characteristics and properties of various electoral systems, and yet, as I have shown here with regard

[28] The relative lack of appreciation of quantitative advances in the study of election results produced by geographers means that, apart from Gudgin and Taylor's (1979) work referred to here several other innovations have had little impact—either within geography or outwith the discipline: Archer and Taylor's (1981) adaptations of factor analysis for analysing normal and critical elections, for example; Rossiter and Johnston's (1981) computer simulations of redistribution in the UK (though see Cirincione et al., 1999); Johnston and Hay's (1982) adaptation of entropy-maximising for tackling the ecological inference problem in *nxmxk* matrices where all of the internal cells are unknown; Jones, Johnston and Pattie, (1992) multilevel modelling of voting patterns—and the adaptations of Brookes's algebra employed here! Separately and together these have allowed the exploration of a number of aspects of electoral geography, such as the volume of tactical voting in the UK (Johnston & Pattie, 1991) and the impact on that of local campaign intensity (Johnston et al, 1997).

[29] And, as the arguments developed by Gudgin and Taylor (1979) and in this lecture show, although an odd shape may indicate a gerrymander, it is neither a necessary nor a sufficient criterion: on shape and gerrymanders, see Monmonier (2001).

to one system only, we know very little—and then for only a small number of countries which use that system.[30]

'Geography matters' is a frequently-employed slogan within our discipline. Studies such as that reported here clearly demonstrate that truth. But more are needed, as are the skills which enable us to convince people of our case. I have offered one approach here, which hopefully will be taken up, evaluated further and then either applied widely or replaced by something better. Then we will be well-placed to show just how important geography is to a central component of most societies—the election of representatives and the creation of governments.

Acknowledgements

I am grateful to John O'Loughlin for inviting me to present the 2001 lecture, and to Charles Pattie, Dave Rossiter and Danny Dorling for their companionship and collaboration over the last decade—which included commenting on a draft of the lecture.

References

Archer, J. C., & Taylor, P. J. (1981). *Section and party: a political geography of American Presidential elections from Andrew Jackson to Ronald Reagan*. Chichester: John Wiley.

Boundary Commission for England (2000). *The review of Parliamentary constituencies in England*. London: Boundary Commission for England.

Brookes, R. H. (1959). Electoral distortion in New Zealand. *Australian Journal of Politics and History*, *5*, 218–223.

Brookes, R. H. (1960). The analysis of distorted representation in two-party single member systems. *Political Science*, *12*, 158–167.

Cirincione, C., Darling, T. A., & O'Rourke, T. G. (2000). Assessing South Carolina's Congressional districting. *Political Geography*, *19*, 189–212.

Denver, D. T., & Hands, G. (1997). *Modern constituency electioneering: local campaigning in the 1992 general election*. London: Frank Cass.

Dyer, M. (2001). The evolution of the centre-right and the state of Scottish Conservatism. *Political Studies*, *49*, 30–50.

Gelman, A., & King, G. (1994). A unified method of evaluating electoral systems and redistricting plans. *American Journal of Political Science*, *38*, 514–554.

Grofman, B. (1983). Measures of bias and proportionality in seats–votes relationships. *Political Methodology*, *9*, 295–327.

Grofman, B., Koetzle, W., & Brunell, T. (1997). An integrated perspective on the three potential sources of partisan bias: malapportionment, turnout differences, and the geographic distribution of party vote shares. *Electoral Studies*, *16*, 454–470.

Gudgin, G., & Taylor, P. J. (1979). *Seats, votes and the spatial organization of elections*. London: Pion.

Hazan, R.Y. (1999). Constituency interests without constituences. *Political Geography*, *18*, 791-812.

[30] Apart from the UK, we have only applied it in New Zealand (Johnston, 1976, 1992) and Jamaica (Johnston and Morrissey, 1984).

Jackman, S. (1994). Measuring electoral bias: Australia, 1949–1993. *British Journal of Political Science, 24*, 319–357.

Johnston, R. J. (1976). Spatial structure, plurality systems and electoral bias. *The Canadian Geographer, 20*, 310–328.

Johnston, R. J. (1992). Electoral geography. In M. Holland (Ed.), *Electoral behaviour in New Zealand* (pp. 25–50). Auckland: Oxford University Press.

Johnston, R. J. (2000). On disciplinary history and textbooks: or where has spatial analysis gone. *Australian Geographical Studies, 38*, 125–137.

Johnston, R. J., & Hay, A. M. (1982). On the parameters of uniform swing in single-member constituency electoral systems. *Environment and Planning A, 14*, 61–74.

Johnston, R. J., & Rossiter, D. J. (1982). Constituency building, political representation and electoral bias in urban England. In D. T. Herbert, & R. J. Johnston (Eds.), *Geography and the urban environment* (Vol. 5) (pp. 113–156). Chichester: John Wiley.

Johnston, R. J., & Morrissey, M. P. (1984). The sources of electoral bias in Jamaican electoral results, 1967–1980. *Caribbean Geography, 2*, 175–185.

Johnston, R. J., & Pattie, C. J. (1991). Tactical voting in Great Britain in 1983 and 1987: an alternative approach. *British Journal of Political Science, 21*, 95–108.

Johnston, R. J., & Pattie, C. J. (2000). Ecological inference and entropy-maximizing: an alternative estimation procedure for split-ticket voting. *Political Analysis, 8*, 333–345.

Johnston, R. J., Pattie, C. J., & Allsopp, J. G. (1988). *A nation dividing? Britain's changing electoral map, 1979–1987*. London: Longman.

Johnston, R. J., Pattie, C. J., MacAllister, I., Rossiter, D. J., Dorling, D. F. L., & Tunstall, H. (1997). Spatial variations in voter choice: modelling tactical voting at the 1997 general election in Great Britain. *Geographical & Environmental Modelling, 1*, 153–179.

Johnston, R. J., Rossiter, D. J., & Pattie, C. J. (1999). Integrating and decomposing the sources of partisan bias: Brookes' method and the impact of redistricting in Great Britain. *Electoral Studies, 18*, 367–388 (and 19, 649–650).

Johnston, R. J., Pattie, C. J., & Rossiter, D. J. (2001). He lost…but he won! Electoral bias and George W Bush's victory in the US 2000 Presidential election. *Representation, 38*, 151–8.

Johnston, R. J., Pattie, C. J., Dorling, D. F. L., & Rossiter, D. J. (2001). *From votes to seats: the operation of the UK electoral system since 1945*. Manchester: Manchester University Press.

Jones, K., Johnston, R. J., & Pattie, C. J. (1992). People, places and regions: exploring the use of multilevel modelling in the analysis of electoral data. *British Journal of Political Science, 22*, 343–380.

King, G. (1989). Representation through legislative redistricting: a stochastic model. *American Journal of Political Science, 33*, 787–824.

King, G., & Browning, R. X. (1987). Democratic representation and partisan bias in Congressional elections. *American Political Science Review, 81*, 1251–1273.

Lennertz, J. (2000). Back in their proper place: racial gerrymandering in Georgia. *Political Geography, 19*, 163–188.

McLean, I. (1995). Are Scotland and Wales over-represented in the House of Commons? *The Political Quarterly, 66*, 250–268.

Monmonier, M. S. (2001). *Bushmanders and bullwinkles: how politicians manipulate electronic maps and census data to win elections*. Chicago: University of Chicago Press.

Morrill, R. L. (1973). Ideal and reality in reapportionment. *Annals of the Association of American Geographers, 63*, 463–477.

Mortimore, R. (1992). *The constituency structure and the Boundary Commission; The rules for the redistribution of seats and their effect on the British electoral system 19501987*. D.Phil thesis, University of Oxford.

Pattie, C. J., Johnston, R. J., & Fieldhouse, E. A. (1995). Winning the local vote: the effectiveness of constituency campaign spending in Great Britain. *American Political Science Review, 89*, 969–983.

Rawlings, H. F. (1988). *Law and the electoral process*. London: Sweet & Maxwell.

Rossiter, D. J., & Johnston, R. J. (1981). Program GROUP: the identification of all possible solutions to a constituency-delimitation problem. *Environment and Planning A, 13*, 231–238.

Rossiter, D. J., Johnston, R. J., & Pattie, C. J. (1997a). The evolution and partisan impact of Scottish

and Welsh over-representation in the redrawing of British Parliamentary constituencies. *Regional and Federal Studies*, *7*, 49–65.

Rossiter, D. J., Johnston, R. J., & Pattie, C. J. (1997b). Estimating the partisan impact of redistricting in Great Britain. *British Journal of Political Science*, *27*, 319–331.

Rossiter, D. J., Johnston, R. J., & Pattie, C. J. (1998). The partisan impacts of non-partisan redistricting: Northern Ireland, 1993–1995. *Transactions, Institute of British Geographers, NS24*, 455–480.

Rossiter, D. J., Johnston, R. J., & Pattie, C. J. (1999). *The Boundary Commissions: redrawing the UK's map of Parliamentary constituencies*. Manchester: Manchester University Press.

Rush, M. E. (2000). Redistricting and partisan fluidity: do we really know a gerrymander when we see one? *Political Geography*, *19*, 249–260.

Tufte, E. R. (1973). The relationship between seats and votes in two-party systems. *American Political Science Review*, *67*, 540–554.

Whiteley, P., & Seyd, P. (1994). Local party campaigning and electoral mobilisation in Britain. *Journal of Politics*, *56*, 242–252.