B.J.Pol.S. **22**, 343–380
Printed in Great Britain

People, Places and Regions: Exploring the Use of Multi-Level Modelling in the Analysis of Electoral Data

K. JONES, R. J. JOHNSTON AND C. J. PATTIE*

There has been considerable recent debate about the importance of local context as an influence on political attitudes and voting behaviour in Great Britain. Resolution of that debate has been difficult, because analytical methods have not been available with which to evaluate the relative importance of both individual voter characteristics and the characteristics of their milieux as independent correlates of attitudes and behaviour. The technique of multi-level modelling has been developed by educational researchers to do just that. It is introduced here and illustrated using data for the 1987 British general election. The preliminary results suggest that place clearly does matter as a component of the processes that influence voters' choices.

Analysts of voting behaviour in Great Britain have sought explanations for variations in the support for a given party at a range of spatial scales, from the individual up to a major region, such as Scotland. Most now recognize that understanding voting patterns calls for sophisticated analysis at a variety of such scales, which involves the conjoint use of both individual and ecological data. Few recognize the methodological difficulties involved in such analysis, however, and the most commonly used procedures adopted ignore some of them (such as the spatial autocorrelation problem) and impose a structure on their analysis that is not necessarily consistent with intuitive modelling of the processes involved.

Recent developments in the statistical modelling of multi-level (or multi-scale) phenomena provide an opportunity to evaluate some of the intuitive models more exactingly than before. This article reports on an initial exploration of the utility of the procedures for developing an understanding of British voting behaviour in recent decades.

VOTING AND SPATIAL SCALE: AN INTRODUCTION

Three major spatial scales relevant to the study of British voting behaviour are identified in this article's title – although several can themselves be divided into component sub-scales and others could be suggested. Discussion here focuses on the three, as does the later empirical illustration of multi-level modelling.

* Respectively, Department of Geography, Portsmouth Polytechnic; Department of Geography, University of Sheffield; and Department of Geography, University of Nottingham. Dr Jones acknowledges support from a Nuffield Social Science Fellowship; Prof. Johnston and Dr Pattie acknowledge support for the Economic and Social Research Council, which permitted the collation of the data sets used.

The fundamental unit in any study of voting is the *individual elector*. The standard theories of electoral behaviour,[1] developed using the British Election Study data, focus on the individual elector and his or her characteristics. A wide range of characteristics has been used, with the individual's position in the occupational class structure receiving most attention; for example, the probability of an individual voting Labour is believed to vary according to class position – usually indexed by occupation – and the notion of a class cleavage dominates British psephology.

Individual electors are not politically socialized in vacuums, however, but learn their political attitudes in a variety of contexts. The most important of these, according to Rose and McAllister,[2] are the *characteristics of the family* in which they were raised – notably their parents' political attitudes and voting predispositions. Since all such studies of voting behaviour look at individuals rather than at families or households, these characteristics are usually incorporated in the operational models as aspects of the individual.

Outside the family, political socialization is influenced by other characteristics of the voters' economic and social positions – such as the sector in which they work (public or private),[3] whether or not they belong to a trade union, and their housing tenure.[4] Union membership, housing, class position and employment sector all act, some more formally than others, as milieux of socialization. Again, however, because most studies of voting rely on data collected from randomly sampled individuals, these are treated as characteristics of the individual. Thus, for example, the probability of voting Labour is considered to vary not only with an individual's occupational class position but also with whether he or she comes from a family with pro-Labour dispositions, lives in a council house or flat, works in the public sector and belongs to a trade union. Together these comprise the main *individual-level influences* on political attitude formation and voting behaviour that have been identified.

Much political socialization occurs in particular spatial contexts, and recent writing on the geography of voting behaviour in Great Britain has emphasized their apparent importance. Attempts to postdict voting patterns using individual characteristics alone have largely failed – people with comparable characteristics are much more likely to vote Labour in some areas than are similar people in other areas, for example.[5] Close attention is thus being paid to the spatial contexts within which socialization takes place, and which form

[1] Notably: D. Butler and D. Stokes, *Political Change in Britain* (London: Macmillan, 1969); B. Särlvik and I. Crewe, *Decade of Dealignment* (Cambridge: Cambridge University Press, 1973); A. Heath, R. Jowell and J. Curtice, *How Britain Votes* (Oxford: Pergamon Press, 1985); and A. Heath, R. Jowell and J. Curtice, *Understanding Political Change* (Oxford: Pergamon Press, 1991).

[2] R. Rose and I. McAllister, *The Loyalties of Voters* (London: Sage Publications, 1990).

[3] See P. Dunleavy, 'The Political Implications of Sectoral Cleavages and the Growth of State Employment', *Political Studies*, 28 (1980), 364–83 and 527–49.

[4] See R. J. Johnston, 'A Note on Housing Tenure and Voting in Britain, 1983', *Housing Studies*, 2 (1987), 112–21.

[5] R. J. Johnston, C. J. Pattie and J. G. Allsopp, *A Nation Dividing?* (London: Longman, 1988).

the arena for political mobilization. We focus here on two particular spatial scales – which we term place and region – while realizing that for most voters there are many such scales, operative at different stages of their political socialization.[6] Much of the early socialization of voters takes place not only in the home but also in its surrounding neighbourhood in which both formal institutions (such as schools) and informal social networks (such as friends and kin) act as learning milieux. Through these, people learn a wide range of attitudes which predispose, but do not necessarily determine, both their interpretation of the structure of society and their position within it. This learning influences not only their future careers but also their images of the society they inhabit. A major consequence of these influences, directly and indirectly, is their impact on where people live – either where they choose to live or where they are constrained to live. In a society that is strongly segregated spatially by occupational class, the majority of people live in neighbourhoods occupied by others similar to themselves. In this way, residential milieu – or *place* – is a continually self-reinforcing context for political socialization.

Two places may be especially important in the socialization process: the place where an individual was raised (especially during his or her adolescent years when political attitudes are first developed) and the place where the individual currently lives. Others may be important too (such as the place where many late teenagers studied and, possibly, first voted in a general election), but they are more difficult to build into models of the socialization process. Unfortunately, few data sets contain sufficient information about the first of these places.[7] Increasingly, however, it is possible to merge data on individuals with information on the neighbourhood where they currently live – usually obtained from census sources.

Although some studies of the so-called neighbourhood effect merely use nominal categories (i.e. working-class or middle-class), others consider the relationship between voting behaviour and place to be a continuous one. Thus, for example, it may be hypothesized that the probability of a working-class individual (however defined) voting Labour increases the greater the proportion of working-class people in that individual's local milieu.

But what is the scale of that milieu; what defines the place in which somebody lives and is socialized? Clearly there is no one such place but several for most individuals. Fully-specified models would incorporate all of them – which would introduce substantial analytical problems because of multi-collinearity. Such a fully-specified model is almost impossible to construct, however, because

[6] An important third relevant spatial scale within Great Britain is country – England, Scotland and Wales. The latter two have particular political parties contesting all constituencies, along with the three that contest all, or virtually all, in Great Britain. For technical reasons, however, we are currently unable to incorporate this further scale in our multi-level modelling work, as current software can only estimate complex models (with random slopes) with no more than three levels.

[7] An exception is G. C. Wright Jr, 'Contextual Models of Electoral Behavior: The Southern Wallace Vote', *American Political Science Review*, 71 (1977), 497–508.

of data constraints. Census data are available at several scales, and developments in the use of Geographical Information Systems increasingly allow them all to be accessed. At present, however, most analyses use data at one scale only – and this is usually not the scale (the polling district) at which individuals are sampled, which has consequences for the statistical analyses.

Place as defined in the previous paragraphs relates to the local context within which individuals live out their daily lives. That is encapsulated within other places that also have an impact on the individual, but which are larger in scale. For the present discussion, we call these *regions*. Much has been written in recent years about the growing regional polarization of the British electorate, to which Curtice and Steed have made notable contributions.[8]

Attempts have been made recently to link this growing regional polarization to the debates within voting studies on both class dealignment and what is sometimes termed pocket-book voting.[9] As regards dealignment, it is generally agreed that the permanent links between classes and parties have been dissolving over recent decades, so that the electorate is more volatile. Rather than develop permanent commitments to particular parties an increasing proportion of voters has been employing what are termed egocentric and sociotropic voting strategies. Egocentric strategies refer to their evaluation of past and potential future government performance in the light of perceived personal situations – usually, although not exclusively, in terms of personal incomes and wealth, which are linked to employment in the vast majority of cases. Sociotropic strategies refer to voter decisions based on evaluations of past and potential future government performance with regard to society as a whole.

Recent decades have seen substantial changes in the pattern of economic prosperity and individual life chances among the various regions of Great Britain. This uneven development should have an uneven impact on sociotropic voting decisions, with individuals, irrespective of their personal characteristics and situations, in regions suffering relative deprivation being less ready than electors in the relatively prosperous regions to reward the incumbent government with votes. The growth of dealignment can thus be linked to the increased regional polarization.

There are difficulties translating this general argument into operational procedures. What regional definitions should be used? Which are relevant to individuals making sociotropic voting decisions? Two types of region have been

[8] J. Curtice and M. Steed, 'Electoral Choice and the Production of Government', *British Journal of Political Science*, 12 (1982), 249–98. See also M. Steed, 'The Core–Periphery Dimension of British Politics', *Political Geography Quarterly*, 5 (1986), S91–S103.

[9] C. J. Pattie, R. J. Johnston, E. Fieldhouse and A. T. Russell, 'A Widening Regional Cleavage in British Voting Behaviour, 1964–1987: Preliminary Explorations', in I. Crewe, P. Norris, D. Broughton and D. Denver, eds, *British Parties and Elections Yearbook 1991* (Brighton, Sussex: Harvester-Wheatsheaf, 1991), pp. 121–44; R. J. Johnston, C. J. Pattie and A. T. Russell, 'Dealignment, Spatial Polarisation and Economic Voting: An Exploration of Recent Trends in British Voting Behaviour,' *European Journal of Political Research*, 20 (1992), forthcoming.

studied:[10] *geographical regions* are contiguous divisions of the national territory, many of which stimulate some degree of personal identity; *functional regions* are groupings of like areas on pre-defined characteristics, irrespective of their geographical location. Both may be relevant to voting decisions, and are therefore incorporated into such studies. Residents of particular geographical regions may develop some common identify predisposing them to vote in certain ways. People in Merseyside, for example, may hear claims in the media that the national economy is doing well, but these claims contradict their local experience and lead them to vote against the incumbent government. Similarly, residents of inner-city council estates may have a common experience of the impact of government policies, whether they live in Sheffield, Swindon or Southwark, and they too may be predisposed to vote in the same way, irrespective of other characteristics.

Both types of region can therefore be incorporated into a model of voting behaviour. As with place, it is possible to introduce these as nominal variables only, proposing (say) that the probability of voting Labour – all other things being equal – is greater in Merseyside or South Yorkshire than in London's outer suburbs. But it is also possible to introduce continuous variables representing, for example, regional variations in unemployment rates or house prices.

Finally, within Great Britain there is the national scale, which divides the island into three *countries* – England, Scotland and Wales. The last two have separate cultural foundations to England,[11] which have been mobilised electorally in recent years by the Scottish National Party and Plaid Cymru respectively. The presence of this further dimension of electoral choice in those countries may result in country-specific differential effects on other parties so that, for example, the probability of an individual voting Conservative in Scotland, all other things being equal, is less than in England.

MULTI-LEVEL MODELLING

We have identified four spatial scales that are relevant to the study of British voting behaviour, therefore – *individual, place, region, and country* – although we are currently unable to operationalize four and so have omitted the last from the illustrative example in the second half of this article. It may be possible to suggest – as Rose and McAllister do[12] – a chronological sequence by which these influence individual voting decisions. Region may come last, for example, with individual characteristics and place coming first and second. In any study of a cross-section of voters, however, our interest is to separate out their relative impact at that one time. A number of different procedures has been explored.[13] Here we explore the potential of *multi-level modelling*, which is statistically

[10] Johnston, Pattie and Allsopp, *A Nation Dividing?*

[11] See Steed, 'The Core–Periphery Dimension'.

[12] Rose and McAllister, *The Loyalties of Voters.*

[13] For example, C. J. Pattie and R. J. Johnson, 'Embellishment and Detail?' *Transactions, Institute of British Geographers*, NS15 (1990), 205–26.

and substantively preferable to any of the others, because it deals simultaneously with both individual and ecological data.

Two-Level Models: Continuous Variables

In order to derive some intuitive understanding of the multi-level approach, take a two-level model, with individuals at level 1 nested within constituencies, or places, at level 2.[14] (Throughout this article, we use 'level' as synonymous with spatial scale. Thus a two-level model involves analysis at two spatial scales – level 1 and level 2.) Figure 1 portrays a range of possible two-level models for the case of one response variable (probability of voting Labour) and one individual-level (i.e. level 1) predictor variable (age of voter), both measured on a continuous scale. For clarity, in all the plots of Figure 1 the cloud of points representing the individual values for probability and age is not shown; there are ten constituencies (i.e. the ten observations at level 2), which are shown when appropriate (i.e. in Figures 1c, d, e and f). Each plot is of substantive interest, and each is represented by a specific statistical model, which separates systematic, 'fixed' features of the data from random aspects, where 'random' means 'allowed to vary'.[15]

In the simplest possible model

$$p_{ij} = \beta_0 + \varepsilon_{ij} \tag{1}$$

it can be postulated that p_{ij}, the probability of voter i in constituency j voting Labour, is merely a function of the countrywide average support for this party, which is represented by the fixed term, β_0, plus individual random fluctuations represented by ε_{ij}. The latter are termed the level-1 random terms and, making the usual assumptions of normality, zero mean, constant variability and no autocorrelation, their variation can be captured by a single measure, σ_ε^2, the variance or the 'average size' of the random variation. In substantive terms, this model implies that voting for Labour is entirely idiosyncratic and is not related to any predictor variable, such as age, in a systematic way. Figure 1a shows this model as an unchanging probability of voting Labour with age.

By contrast, in a simple bivariate model

$$p_{ij} = \beta_0 + \beta_1 x_{ij} + \varepsilon_{ij} \tag{2}$$

the additional 'slope' term, β_1, measures the change in the probability of voting Labour for a unit change in age, where x_{ij} measures the age of each individual in every constituency.[16] The random terms in this model, ε_{ij}, can again be summarized by σ_ε^2, but now the 'residuals' represent the variation remaining

[14] For a fuller description, see K. Jones, 'Specifying and Estimating Multi-Level Models for Geographical Research', *Transactions, Institute of British Geographers*, NS16 (1991), 148–59; and K. Jones, *Multilevel Models for Geographical Research* (Norwich: Concepts and Techniques in Modern Geography, Environmental Publications, 1991).

[15] In the terminology of multi-level modelling, random means *allowed to vary*, not haphazard.

[16] Throughout this article, age for each individual is measured as that person's deviation from the mean age for the entire sample.

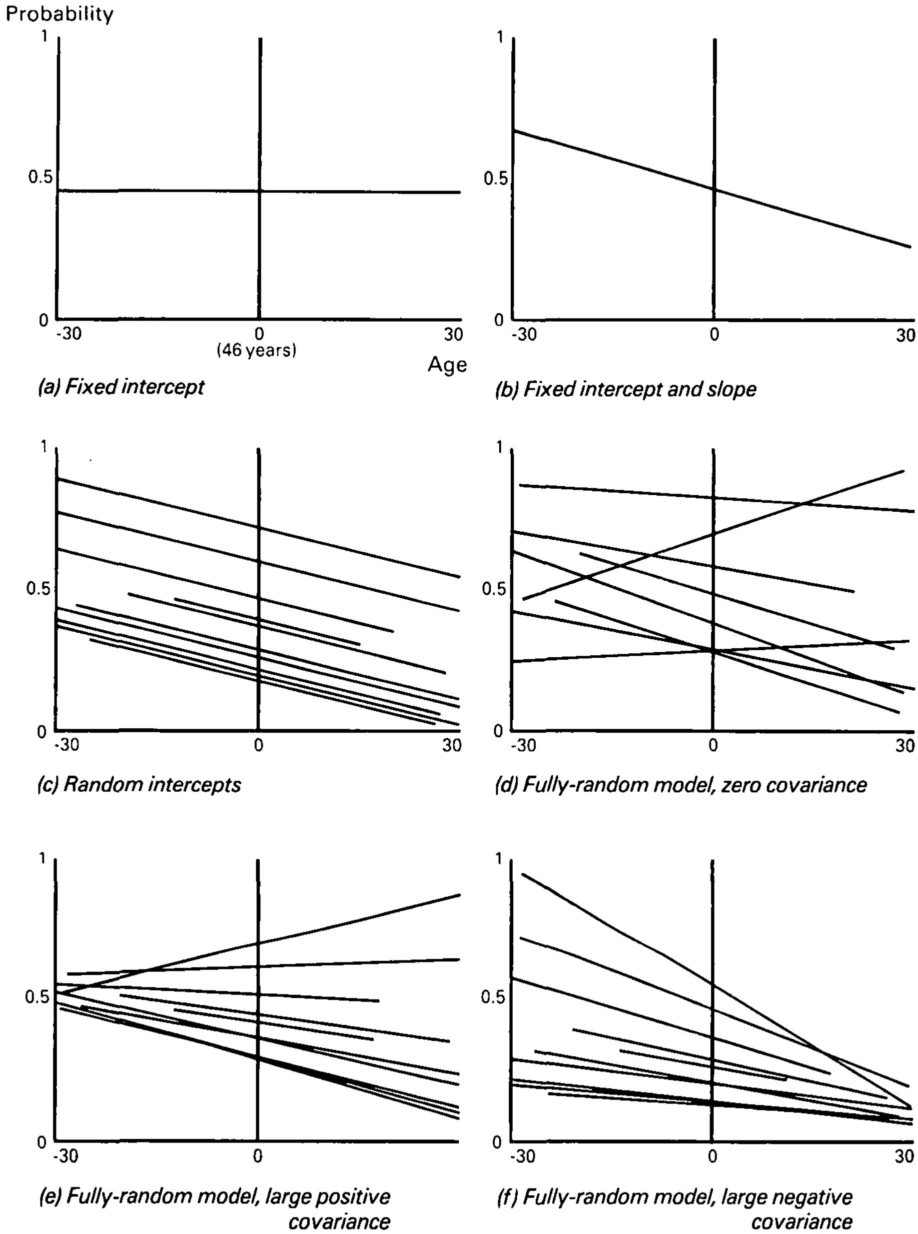


Fig. 1. Alternative models for multi-level modelling with a continuous predictor variable: in (a) and (b) a single relationship between the dependent and predictor variables is found throughout the sample; in (c)–(f) the relationship varies from place to place

after the individual's age has been taken into account. If the predictor variable is centred around its mean, so that x_{ij} is measured as a deviation, the 'intercept' term, β_0, will represent the probability of voting Labour for an individual of average age. Figure 1b portrays this model as a single (negative) relationship between age and voting Labour, which is found unvaryingly across all areas. (Because the age variable has been centred around a mean of 46 years, a value of -20 represents a 26-year old voter.) If this model is appropriate, place does not matter; younger voters everywhere prefer Labour, and the probability of voting Labour declines uniformly with age in every constituency sampled.

In both Equations (1) and (2), as the model is only specified at a single level, the intercept and slope parameters are fixed and unvarying. The remaining plots of Figure 1, however, display two-level models. The first of these is the 'random-intercepts' model of Figure 1c which consists of a set of parallel lines, one for each constituency.[17] These allowed-to-vary intercepts permit different places to have different levels of support for Labour around the overall, fixed average for all constituencies. This is achieved by respecifying the voter-level (level 1), micro-model of Equation (2):

$$p_{ij} = \beta_{0j} + \beta_1 x_{ij} + \varepsilon_{ij} \tag{3}$$

so as to allow the intercept terms, β_{0j}, to vary in a macro-model:

$$\beta_{0j} = \beta_0 + \mu_j \tag{4}$$

The μ_j terms are the constituency-level random terms at level 2; making the same assumptions as for the level-1 random terms, they can be summarized by a single variance term, σ_μ^2. In effect, such a formulation allows for observations within constituencies to be non-independent, that is autocorrelated with voters in a constituency being more alike than a random sample.[18] Examining Figure 1c, we can see the same slope for age in each constituency, but the random intercepts allow the probability of voting Labour to be consistently higher in some places than others at every age.[19]

The remaining parts of Figure 1 display models in which both slopes and intercepts for each constituency are allowed to vary around the overall fixed

[17] See fn. 15 above.

[18] This is an aspect of multi-level modelling that makes it particularly valuable for the analysis of British Election Study (BES) data. In the surveys conducted for the BES, the respondents are selected according to cluster sampling designs (as described in Butler and Stokes, *Political Change in Britain*, and Heath *et al.*, *How Britain Votes*). This introduces spatial autocorrelation: people living in the same area within a constituency – in this case, a polling district, which is the sampling frame used – are more likely to be similar in their socio-economic characteristics, for example, than are people selected randomly from within the entire constituency. This introduces imprecision to the estimation of relationships at the individual level (between age and voting, for example): by explicitly including the constituencies as a higher level, multi-level modelling automatically overcomes the problems of mis-estimated precision that are inherent in a clustered, multi-stage sampling design.

[19] The lines on the graph (a) in Figure 1c are drawn to extend from the minimum to the maximum age in each constituency, for it is good practice not to 'predict' beyond the range of the available data.

averages for all constituencies. Potentially, this allows differing relationships between age and Labour voting in different constituencies. There may be constituencies where the elderly are relatively pro-Labour, others where they are relatively anti-Labour, and yet others where there is no relationship between age and Labour voting. Such a fully random two-level model is specified by allowing the slope term of the micro-model:

$$p_{ij} = \beta_{0j} + \beta_{1j}x_{ij} + \varepsilon_{ij} \tag{5}$$

to vary according to a further macro-model:

$$\beta_{1j} = \beta_1 + \Gamma_j. \tag{6}$$

The Γ_j are another set of level-2 random terms and, making the usual assumptions, they can be summarized in a single variance term, σ^2_Γ. The different forms of this fully-random model that are shown in Figures 1d, e and f are produced by different values of the covariance term, $\sigma_{\mu\Gamma}$. This term allows the slopes and intercepts to be 'correlated', so that with a positive covariance, constituencies with high overall support for Labour have steep positive slopes for age (e). In contrast, a negative covariance implies that where the support for Labour is high there is a strong negative relationship with age (f). If the covariance is zero, the level of constituency support for Labour tells us nothing about the age/support relationship (d). Another way of looking at these different relationships is shown clearly on the plots: the positive covariance in (e) implies that for the young voter, place does not make much of a difference; the negative covariance of (f) suggests that place matters for the young but not for the elderly; while the zero covariance of (d) means that there is a wide variety of differential support in every age group.

To summarize, putting equations (4), (5) and (6) together gives the fully random population model:

$$p_{ij} = \beta_0 + \beta_1 x_{ij} + (\Gamma_j x_{ij} + \mu_j + \varepsilon_{ij}) \tag{7}$$

in which the probability (p_{ij}) of individual i in constituency j voting Labour is specified as a function of: β_0, which is the countrywide preference for Labour; $\beta_1 x_{ij}$, which is the countrywide relationship between age and Labour preference; $\Gamma_j x_{ij}$, which is the differential constituency-specific relationship between age and Labour preference; μ_j, which is the residual differential constituency preference for Labour; and ε_{ij}, which is the residual individual idiosyncratic difference in preference for Labour.

When this model is calibrated on sample data, if the estimate, σ^2_μ, is found to be small in relation to sampling error, there is no tendency for constituencies to vary in their support for Labour once age is taken into account; if σ^2_Γ is small the relationship between age and Labour support is consistent across the country; if $\sigma_{\mu\Gamma}$ is small, then variations in the support–age relationships are unrelated to constituency average support. If all three terms are small, multi-level models are not needed and place in the contextual sense does not matter; there is no geography of voting. Any apparent differences between places if β_1 is 'significant' are really an outcome of variations in the age profiles of the constituencies. Clearly, multi-level models represent a powerful quantitative approach to exploring empirically the way in which places matter.

Two-Level Models: Categorical Predictors

The discussion has so far been based on a predictor variable that is continuous, but it is much more common for theories of voting, as the term cleavage implies, to use qualitative rather than quantitative descriptions of individual voters. Such categorical states are included in the model by sets of dummy variables. Thus, in the case of a consumption cleavage based on tenure in which there are three categories ('owner-occupier', 'local authority tenant' and 'other'), there would need to be two dummy variables in the micro-model:

$$p_{ij} = \beta_0 + \beta_1 x_{1ij} + \beta_2 x_{2ij} + \beta_3 x_{3ij} + \varepsilon_{ij} \tag{8}$$

If x_2 is coded 1 for local authority, 0 otherwise, and x_3 is coded 1 for other, 0 otherwise, the overall intercept, β_0, is the probability of an individual of average age who is an owner-occupier voting Labour. The fixed terms β_2 and β_3 will represent the differential probability of voting Labour, irrespective of age, if the voter lives in the local authority or other tenure categories, respectively. Such a formulation presumes that place makes no difference, and the essence of such a model is conveyed graphically in Figure 2a. However, if the parameters of the micro-model are indexed for each j constituency

$$p_{ij} = \beta_{0j} + \beta_{1j} x_{1ij} + \beta_{2j} x_{2ij} + \beta_{3j} x_{3ij} + \varepsilon_{ij} \tag{9}$$

macro-level equations can be specified so that the effects for each category can potentially vary from constituency to constituency. Thus, in a random intercepts model:

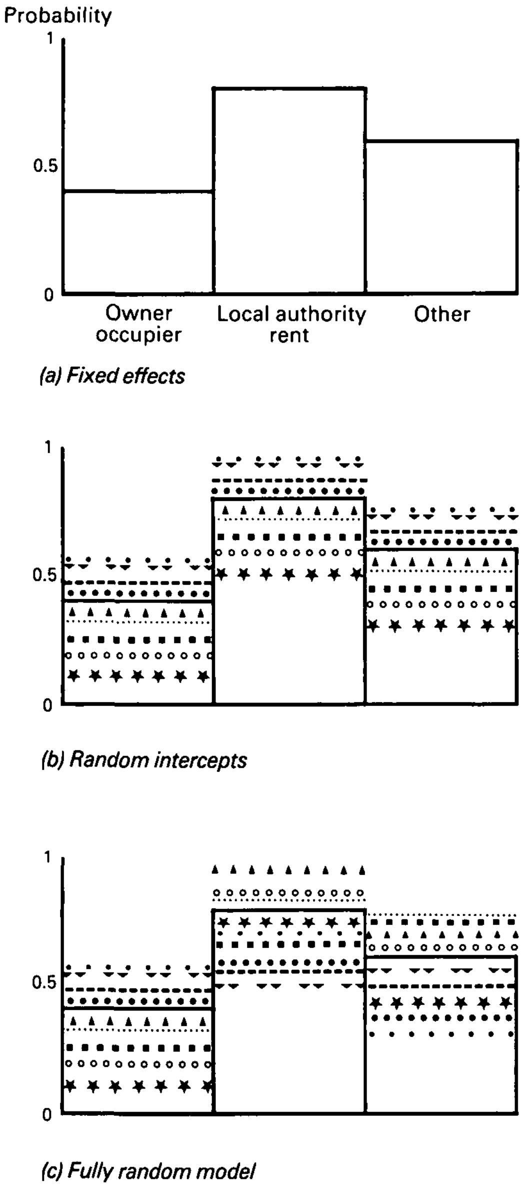
$$\beta_{0j} = \beta_0 + \mu_j \tag{10}$$

there will be a single overall constituency differential which will apply, as already discussed, to age (Figure 1c) and additionally to all three tenure categories (Figure 2b). Moreover, a fully-random model can be specified by additionally including macro-models for each predictor variable

$$\begin{aligned} \beta_{1j} &= \beta_1 + \Gamma_{1j} \\ \beta_{2j} &= \beta_2 + \Gamma_{2j} \\ \beta_{3j} &= \beta_3 + \Gamma_{3j} \end{aligned} \tag{11}$$

These equations allow for a different differential effect for each tenure category for each constituency, as is shown in Figure 2c. The overall 'patterning' of the results depends not only on the size of the level-2 variances but also on the covariances between the level-2 random effects.

Conceptually, it is very easy to postulate complex models but data requirements must be borne in mind, for the fully-random model resulting from the



Equations

(a)

Owner occupier $= B_0$

Local authority rent $= B_0 + B_2$

Other $= B_0 + B_3$

(b)

Owner occupier $= B_0 + \mu_j$

Local authority rent $= B_0 + B_2 + \mu_j$

Other $= B_0 + B_3 + \mu_j$

(c)

Owner occupier $= B_0 + \mu_j$

Local authority rent $= B_0 + B_2 + \mu_j + \Gamma_{2j}$

Other $= B_0 + B_3 + \mu_j + \Gamma_{3j}$

Fig. 2. Alternative models for multi-level modelling with a three-category predictor variable: in (b) and (c), in contrast to (a), the relationship between the two variables is found to vary from place to place

combination of Equations (9), (10) and (11) requires the estimation of ten level-2 parameters – that is, of four variances and six covariances.

Unless there is a considerable number of constituencies, these parameters will be poorly estimated, and it is unlikely that 'significant' level-2 effects will be detected. Although there is little research on the optimal design for multi-level models, Paterson and Goldstein suggest that twenty-five individuals in each of twenty-five higher-level units is a desirable minimum, and they prefer 100 higher-level units.[20]

Two-Level Models: Binary Response

The observed response in the case of voting is binary, not continuous. While this can be modelled by treating the response as if it were the probability of voting Labour, this formulation results in three problems: (1) nonsensical values, outside the range 0.0–1.0; (2) the introduction of an inappropriate non-linear form; and (3) variance heteroscedasticity.[21]

These problems are well-known in single-level modelling, where they are overcome in the framework of generalized linear models (GLM) by the specification of a logit link function between the response and predictor variables, and a binomial distribution for the random part:[22] the GLM framework has recently been extended to include multi-level models.[23] A fully random, two-level model with three predictors and a binary response can be postulated with the same macro-terms as in Equations (10) and (11), but with the following micro-equation:

$$E(y_{ij}) = p_{ij}$$

where

$$p_{ij} = \frac{\exp(\beta_{0j} + \beta_{1j}x_{1ij} + \beta_{2j}x_{2ij} + \beta_{3j}x_{3ij})}{1 + \exp(\beta_{0j} + \beta_{1j}x_{1ij} + \beta_{2j}x_{2ij} + \beta_{3j}x_{3ij})}. \tag{12}$$

The y_{ij} are the observed response of 1 or 0 for the ith voter in the jth constituency, and p_{ij} is a theoretical entity, the probability of voting Labour. This model is non-linear, which resolves the problem of appropriate functional form. For estimation it can be linearized by taking a logit transformation:

$$E[\mathrm{Log}_e(p_{ij}/(1 - p_{ij}))] = \beta_{0j} + \beta_{1j}x_{1ij} + \beta_{2j}x_{2ij} + \beta_{3j}x_{3ij}. \tag{13}$$

[20] L. Paterson and H. Goldstein, *New Statistical Methods for Analysing Social Structures: An Introduction to Multilevel Models* (London: Institute of Education, forthcoming). Their advice stems from educational research, and it may be possible to manage with a smaller sample if the contextual effects for constituencies are greater than those for schools.

[21] For fuller details, see N. Wrigley, 'The Use of Percentages in Geographical Research', *Area*, 5 (1973), 183–6.

[22] P. McCullagh and J. A. Nelder, *Generalized Linear Models* (London: Macmillan, 1986).

[23] H. Goldstein, 'Nonlinear Multilevel Models', *Biometrika*, 78 (1991), 45–52; N. Longford, 'Multilevel Models for Data with a Non-Normal Distribution', *Multilevel Modelling Newsletter*, 2 (1990), p. 2.

The entity $p_{ij}/(1-p_{ij})$ represents the odds of an individual voting Labour and ranges from zero to infinity. The natural logarithm of this value, the logit, ranges from minus to plus infinity as p_{ij} ranges from zero to one. As the logit cannot exceed infinity, p_{ij} must be bounded by zero and one, thereby overcoming the problem of nonsensical values. Consequently, for example, the fixed slope parameter β_1 represents the change in the log-odds of voting Labour as the value of age increases by one year.

As with single-level GLMs the variance heterogeneity problem is modelled by assuming a binomial distribution for the voter-level random part. While this variance is always estimated from the data in normal-theory models, it is usual to 'fix' the variance at 1 for the case of the binomial random distribution.[24] Such a procedure may be justified theoretically but be inappropriate in practice, for this would presume that the only variation remaining unexplained at level-1 is due to sampling fluctuations for a response with a discrete outcome. Such an expectation remains a counsel of perfection, for there may be model mis-specification, or extra 'dependency' or 'clustering', which can occur when the level-1 data are spatially organized. Consequently, it seems reasonable to unconstrain the level-1 variance and estimate it from the data on the basis of an 'over-dispersed' distribution with extra-binomial distribution.

Extending the Models

So far we have developed a series of multi-level models by including a number of predictors to represent the characteristics of individual voters, which are then allowed to vary at the higher, constituency, level. Further extensions are achieved by continuing the basic principle of specifying a model at each level. For example, predictor variables can be specified at level 2 so as to account for the constituency-level differences. Thus, in a random-intercepts model, Equation (9) can be extended to include constituency variables:

$$\beta_{0j} = \beta_0 + \alpha_1 w_{1j} + \alpha_2 w_{2j} + \mu_j \tag{14}$$

so that the constituency average support (β_{0j}) for Labour (after taking account of the voter's individual age and tenure) is modelled as a function of national Labour support (β_0) plus, for example, an effect, α_1, for the unemployment rate in the constituency (w_1), plus an effect, α_2, for the percentage of mining workers in the constituency, w_2, plus a random constituency difference, μ_j. These higher level predictors can be 'global' and only occur at that level, or they may be 'aggregate' and be derived as some summary (for, example, the mean or variance) of the level-1 predictors. Another elaboration is to extend the model to higher levels, so that it is possible to have a four-level hierarchy of voter (1) in ward (2) in constituency (3) in region (4) and to specify a model with a mixture of random and fixed terms at each level. The VARCL software is capable of fitting a random-intercepts model with up to nine levels, but

[24] J. Nelder and B. Baker, *GLIM3 Manual* (Oxford: NAG Publications, 1981).

again the requirements of a sufficient number of units at each level for effective estimation must be considered.[25]

Estimation

The basic concepts of multi-level models have been discussed for some time,[26] but it was not until the later 1980s that computationally efficient estimation schemes were developed and implemented in general-purpose software. The empirical results to be discussed later were derived using Goldstein's 'iterative generalized least squares' estimator as implemented in the ML3 interactive package using macros that have been written to deal with the binary response case.[27] Of the four multi-level packages that are currently available, ML3 is particularly flexible at specifying complex models.[28] Moreover, as it is part of a general statistical package, it is straightforward to manipulate data prior to analysis and to produce high-quality graphical output.

While it is not appropriate to deal with the estimation algorithm in any detail, it is important to appreciate that multi-level estimates have the very useful property of being precision-weighted. The fully-random plots of Figure 1d–f suggest that multi-level modelling is nothing more than the fitting of sets of regression lines, one for each constituency. Such a 'separate' estimation strategy is likely to be a poor one with particularly imprecise estimates when there are few individuals in a constituency or when the range of variation in the predictor variable for each area is limited. By contrast, in multi-level models, the random terms are conceived not as independent estimates but as coming from a distribution. It is the variance of the distribution that is estimated, and consequently all of the information is effectively 'pooled' in the estimation of the difference for each constituency.[29] Taking the case of the estimation of the random slopes, the fixed slope, β_1, is a weighted average of all the constituency slopes, $\beta_1 + \Gamma_j$, where greater weight is given to the most reliably estimated ones. Moreover, the constituency-level random slopes will be shrunk towards the fixed overall slope if they are unreliably estimated. However, if there is a precisely estimated place-specific slope, it is substantially immune from the influence of the relations that exist in other places. The recognition of the hierarchies and autocorrelation in the data results in improved estimation whereby the estimates 'borrow strength' from wherever information is to be found in the data.

[25] N. Longford, *Manual for VARCL* (Princeton, NJ: Education Testing Services, 1988).

[26] L. Burstein, 'Issues in the Aggregation of Data', in D. C. Berliner, ed., *Review of Research in Education* (Washington, DC: American Educational Research Association, 1980).

[27] H. Goldstein, *Multilevel Models in Educational and Social Research* (London: Charles Griffin, 1987); R. Prosser, J. Rasbash and H. Goldstein, *ML3: Software for Three-Level Analysis* (London: Institute of Education, 1991).

[28] I. G. G. Kreft, J. De Leeuw and K-S. Kim, *Comparing Four Different Packages for Hierarchical Linear Regression* (Los Angeles: UCLA Statistical Series No. 50, 1990).

[29] K. Jones and G. Moon, 'Medical Geography', *Progress in Human Geography*, 15 (1991), 437–43.

AN ILLUSTRATION

To illustrate the use of multi-level modelling in the study of British voting behaviour, we have conducted initial analyses of the Labour vote among respondents to the 1987 British Election Study survey.

Data Collation

Data from the 1987 British Election Study were arranged to allow modelling on three levels: the individual voter;[30] the constituency; and the geographical region (we employed the regionalization used by *The Economist*).[31] Our dependent (or response) variable was voting choice in 1987. In the survey, as in the polling booth, this involved five main choices (the three main parties, the nationalist parties in Scotland and Wales, and 'Did Not Vote'), as well as several minor party options. Discounting the minor parties, there are therefore ten sets of binary choices (Labour or Conservative; Labour or Alliance; Labour or Nationalist; and so on).[32] To simplify the situation, we analysed only those respondents who voted either Conservative or Labour: as explained below, it is possible to carry out multi-level analyses of polytomous dependent variables, but at the cost of 'consuming' an extra level so that it would not be possible given current software to analyse complex models with the three levels of individual, constituency and region.

Our analyses are therefore based on 2,281 individual respondents, living in 250 of Britain's 633 constituencies (an average of nine voters in each constituency), with the constituencies themselves divided between twenty-two regions (on average, eleven constituencies per region). The data are not evenly balanced between regions and constituencies, however (Table 1): some regions, particularly in the south of the country, are heavily represented in the sample, while others (particularly Devon and Cornwall) contain only a few respondents and sampled constituencies. Broadly, this is what we would expect, given that the south of the country is the most populous region and contains the greatest number of parliamentary seats.

This may lead to problems, however. As we have already seen, Paterson and Goldstein recommend a data structure with at least twenty-five respondents in each of twenty-five higher-level units as a minimum.[33] Our data fall short of this. There are relatively few respondents in each constituency, which may cause difficulties in estimating constituency effects: for instance, if the sample for a particular constituency is drawn from an atypical polling district, then estimates for that constituency are likely to be 'unusual'. However, if constituency effects are stronger than the school effects that Paterson and Goldstein

[30] For details of the survey, see Heath *et al.*, *Understanding Political Change*.

[31] Johnston, Pattie and Allsopp, *A Nation Dividing?*

[32] C. B. Begg and R. Gray, 'Calculation of Polychotomous Logistic Regression Parameters using Individualised Regressions', *Biometrika*, 71 (1984), 11–18.

[33] Paterson and Goldstein, *New Statistical Methods*.

TABLE 1 *Data Structure*

Level	Number	Meaning	
1	2,281	Individual voters	
2	250	Constituencies (i.e. polling districts)	
3	22	Economic regions	
	Economic Region	**Constituencies**	**Voters**
1	Strathclyde	10	89
2	East-Central Scotland	8	66
3	Rural Scotland	5	32
4	Rural North	12	123
5	Industrial North-East	11	125
6	Merseyside	7	72
7	Greater Manchester	9	89
8	Rest of Northwest	13	134
9	West Yorkshire	8	77
10	South Yorkshire	6	70
11	Rural Wales	4	38
12	Industrial South Wales	9	95
13	West Midland Conurbation	11	101
14	Rest of West Midlands	13	100
15	East Midlands	17	159
16	East Anglia	9	81
17	Devon and Cornwall	7	63
18	Wessex	13	133
19	Inner London	10	71
20	Outer London	20	160
21	Outer Metropolitan	29	253
22	Outer South-East	19	159

are interested in, this need not be too great a difficulty. Furthermore, regional effects are estimated on only twenty-two observations,[34] with consequent difficulties in estimating significance levels. To a large extent, however, we have had to accept these limitations, since problems arising from sample design were effectively beyond our control: moreover, increasing the number of regions being studied would have reduced the number of voters in each region, with obvious consequences for the overall reliability of our estimates.

In order to model respondents' voting behaviour, we defined a number of predictor variables, drawn from the survey, on respondents' social class, on whether they were in employment, on their housing tenure and on their age (Table 2). We employed the class schema adopted by Heath *et al.*,[35] which divides the electorate into: the salariat (professional and managerial grades);

[34] The places are indeed the observations, and the greater the number the better. Having many voters provides information on the relationships between response and predictor variables *within* a place, but many places are needed to estimate the differences *between* places.

[35] Heath *et al.*, *How Britain Votes*.

TABLE 2 *Variables Used in the Analysis*

Level	Label	Variable	%	Mean
1		Voted Labour	40	
1		Voted Conservative	60	
1		Working-class	38	
1	PRISAL	Private-sector salariat	15	
1	NONMAN	Routine non-manual	25	
1	BOURG	Petty bourgeoisie	8	
1	FORSUP	Foreman/supervisors	5	
1	PUBSAL	Public-sector salariat	9	
1		Employed	95	
1	UNEMP	Unemployed	5	
1		Owner-occupier	71	
1	COUNC	Local-authority tenant	21	
1	PRITEN	Other	8	
1	AGE	Age		46
2	ConEmp	Percent unemployed 1986		12
2	ConCha	Percent change in unemployment		−32
2	ConMine	Percent workforce in mining		3

the petty bourgeoisie (small farmers, the self-employed, etc.); routine non-manual workers (white-collar staff); foremen and supervisors; and the working class. We refined this schema by dividing the salariat into those who work in the private sector and those who work in the public sector, so introducing the possibility of sectoral cleavages.[36] As the largest single class, the working class forms our reference group. Housing tenure divisions (we differentiate between owner-occupiers, council tenants, and other renters, with owner-occupiers as the reference group) allow us to operationalize consumption cleavages. These were coded as dummy variables, with a value of 1 when the respondent shared the attribute and 0 when he or she did not. The continuous variable for respondent's age was deviated around the mean for the sample (46 years old), so that for a 46-year old, the variable took a value of 0; for a 20-year old it was −26, and for a 60-year old it was +14. Deviating age around the mean like this has several analytical advantages: (1) it prevents large numbers, which could cause numerical instability, from entering the models; (2) the intercept is within the range of the data and is given a straightforward interpretation; and (3) in random slopes, the positioning of the vertical axis gives very different estimates of intercepts and hence of place-contextual effects (Figure 1). The dummies and the age variables were coded in such a way that the reference category (zeros for each dummy variable and for the age variable) was the stereotypical voter in the sample (i.e. a 46-year old, working-class, owner-occupier in employment).

In addition to individual level (level-1) data from the 1987 BES, our data

[36] Dunleavy, 'The Political Implications'.

set also contains constituency (level-2) data drawn from the 1981 Census Small Area Statistics and from Department of Employment data for the later 1980s (Table 2). These contextual data include the percentage in mining in 1981; the percentage unemployed in 1986; and the change in the percentage unemployed. Research based on ecological regression of constituency election results suggests that these factors have been important predictors of voting behaviour:[37] for instance, Labour does better than expected in constituencies with a large mining work-force, and worse in mainly agricultural areas.

Discussion

The interactive package ML3 was used to calibrate a series of models to reveal different aspects of the voting data. Table 3 provides summary statistics for a number of increasingly complex two-level (voters in constituencies) models. *Model A* is the constant-only model and this merely partitions the variability into two levels around a fixed mean. The fixed estimate (Table 3) represents the log-odds of voting Labour and, when transformed, its value indicates that the percentage of Labour voters in the entire sample (excluding, of course, those who voted for third and minor parties, and those who abstained) is 40 per cent.[38] In this model, the level-1 variance has been constrained to 1 (on the basis of an exact binomial distribution), while the level-2 variance has been estimated from the data. The size of this constituency variance of 0.88 about a fixed average of −0.38 suggests substantial variation. This can be assessed for significance by calculating the ratio of this estimate to its standard error. If the ratio is in excess of plus or minus 2, the estimate is judged significantly different from zero at the 0.05 level.[39] This ratio is shown under the heading Z in Table 3.[40] For *Model A*, the level-2 variance is significant, and the differences between constituencies are more than would be expected from sampling fluctuations. Another way of representing these results is to calculate the autocorrelation coefficient, which measures the similarity of individual voting within constituencies. This is given by the ratio of the level-2 variance to the total random variance (level 1 plus level 2).[41] The resultant value, 0.47, suggests strong patterning and that there is a need to analyse these data in a multi-level model. Table 4 provides a list of the twenty largest positive and negative constituency random terms that form part of the level-2 distribution. When transformed these suggest that in the most pro-Labour place in

[37] R. J. Johnston, C. J. Pattie and L. C. Johnston, 'Great Britain's Changing Electoral Geography', *Tijdschrift voor Economische en Sociale Geografie*, 81 (1990), 189–206.

[38] To avoid confusion, logits will be given in the text with decimal points; when transformed into proportions they are multiplied by 100 and given as percentages.

[39] This rule of thumb is used because the exact distribution theory is not known; the advice is to treat 'marginal' significance with care. See S. W. Raudenbush and A. S. Bryk, 'A Hierarchical Model for Studying School Effects', *Sociology of Education*, 59 (1986), 1–17.

[40] As degrees of freedom increase, Z-ratios equate to t-ratios.

[41] As indicated above, the level-1 variance is set to 1.00, so the autocorrelation coefficient is $0.88/(1+0.88)=0.47$.

TABLE 3 *Summary of Estimated Two-level Models: Individuals in Constituencies**

	Model A Estim.	Z	Model B Estim.	Z	Model C Estim.	Z	Model D Estim.	Z
FIXED								
CON	−0.38	–	−0.20	–	−0.21	–	−0.21	–
UNEMP			0.84	3.6	0.82	3.5	0.77	3.7
AGE			−0.02	−5.9	−0.02	−5.8	−0.02	−5.7
PRISAL			−1.36	−8.3	−1.36	−8.1	−1.31	−8.8
NONMAN			−0.81	−6.4	−0.81	−6.4	−0.79	−6.9
BOURG			−1.24	−6.1	−1.23	−6.0	−1.19	−6.5
FORSUP			−0.32	−1.5	−0.33	−1.6	−0.32	−1.7
PUBSAL			−0.49	−2.8	−0.50	−2.9	−0.47	−3.0
COUNC			1.30	9.8	1.32	10.0	1.26	10.5
PRITEN			0.57	3.2	0.58	3.2	0.56	3.4
RANDOM								
Level 2								
CON	0.88	7.2	0.49	5.2	0.51	5.3	0.60	6.2
AGE/CON					−0.004	−1.3	−0.004	−1.5
AGE					0.0000	0.01	0.0003	1.9
Level 1								
CON	1.00	–	1.00	–	1.00	–	0.79	–

* *Model A*: two-level null model (constant only), assuming binomial, level-1 variance.
Model B: as A, but additionally includes nine individual-level predictors.
Model C: as B, but with a complex random part in which the relationship with age is allowed to vary between constituencies.
Model D: as C, but unconstrained level-1 variance.
All estimates are on a logit scale.

the sample, Renfrew West and Inverclyde, over 81 per cent of the respondents favoured Labour, but this drops to 18 per cent in Beverley, at the other extreme.[42]

[42] In a perfectly designed and executed survey, these forty areas would indeed be the safest seats in the country for either the Labour or Conservative parties. There are few surprises in the list, indicating a generally representative sample, the exceptions being Thurrock, Walsall North, Tynemouth, Nottingham South, Renfrew West and Inverclyde, and Blyth Valley. In the case of Renfrew West and Inverclyde, it may seem odd that this was 'the most pro-Labour place in the sample', when in fact the Conservative party won the seat in 1983. The similar finding in Model B indicates that the place was the most pro-Labour, *given the characteristics of its population*. It should also be noted that the cluster sampling procedure used in the BES may have resulted in the selection of a polling district that was very unrepresentative of the constituency as a whole. This calls for further work which uses the polling district as one of the levels in a multi-level model, an approach which is as yet not possible due to the inability to estimate complex models at more than three levels with existing software.

TABLE 4 *Twenty Largest Effects for Constituencies, Models A and B**

MODEL A		MODEL B	
Constituency	Logit	Constituency	Logit
Beverley	−1.14	Spelthorne	−0.82
Ravensbourne	−1.14	Cambridgeshire South West	−0.82
Saffron Walden	−1.14	Peterborough	−0.75
Spelthorne	−1.14	Gloucestershire West	−0.75
Welwyn Hatfield	−1.14	Beverley	−0.72
Devon West and Torridge	−1.10	Hexham	−0.71
Bexleyheath	−1.06	Walsall North	−0.71
Cheltenham	−1.06	Taunton	−0.69
Daventry	−1.06	Woking	−0.69
Devon North	−1.06	Dorset South	−0.67
Tayside North	−1.06	Thurrock	−0.66
Cirencester and Tewkesbury	−1.03	Bosworth	−0.66
Brentwood and Ongar	−1.00	Cheltenham	−0.65
Cambridgeshire South West	−1.00	Skipton and Ripon	−0.65
Ilford North	−1.00	Tayside North	−0.64
Thurrock	−1.00	Ravensbourne	−0.64
Walsall North	−1.00	Newbury	−0.64
Berkshire East	−1.00	Cirencester and Tewkesbury	−0.64
Tynemouth	−1.00	Bexleyheath	−0.63
Hertfordshire South West	−0.96	Berkshire East	−0.63
Sheffield Brightside	1.20	Hull East	0.70
Sheffield Heeley	1.20	Livingston	0.72
Glasgow Provan	1.23	Newham North West	0.73
Nottingham South	1.23	Sheffield Brightside	0.83
Birmingham Sparkbrook	1.28	Liverpool Walton	0.85
Huddersfield	1.35	Durham North	0.85
Hamilton	1.39	Birkenhead	0.85
Houghton and Washington	1.46	Merthyr Tydfil and Rhymney	0.86
Preston	1.47	Keighley	0.92
Blyth Valley	1.53	Leigh	0.96
Leigh	1.56	St Helens North	0.97
Merthyr Tydfil and Rhymney	1.56	Preston	1.00
Bishop Auckland	1.58	Bishop Auckland	1.05
Blaenau Gwent	1.58	Monklands East	1.05
Hull East	1.63	Huddersfield	1.06
Sheffield Attercliffe	1.63	Gower	1.09
Durham North	1.73	Sheffield Attercliffe	1.20
Monklands East	1.73	Ogmore	1.25
Ogmore	1.78	Renfrew West and Inverclyde	1.25
Renfrew West and Inverclyde	1.85	Blaenau Gwent	1.43

* *Model A*: differential logits for null model that includes only the constant.
Model B: differential logits for model that includes nine individual-level predictors.

Model A suggests that there are major differences between places in their voting behaviour. But this may be an artefact of, for example, the class composition of the constituencies. Thus, Beverley's low vote for Labour may be explained by a low proportion of working-class inhabitants. This hypothesis, that contextual effects are really a result of a place's socio-economic composition, is evaluated in Model B, when a number of voter characteristics are included in the model at level 1. The key to the interpretation of the estimates for this model (Table 3) is to realise that the fixed constant represents the log-odds of voting Labour for the stereotypical voter, who is a 46-year-old employed member of the working class, and is also an owner-occupier. The fixed estimates for the categorical terms represent differentials, so that the log-odds in favour of Labour increase significantly if the voter is unemployed, or is in the council house or other tenure categories. Being a council tenant is most likely to raise the propensity to vote Labour above that of the stereotypical voter, with unemployment also a major contributor. In contrast, the log-odds, and hence the probability of voting Labour rather than Conservative, decrease significantly if the voter is a member of the private-sector salariat, the petty bourgeoisie, the non-manual class, or the public-sector salariat: of those less likely to vote Labour than the stereotypical voter, the public sector salariat was relatively the most pro-Labour and the private sector salariat the most anti-Labour. In this sample, however, foremen and supervisors do not differ significantly from the working class in their preference for Labour. The negative and significant sign for age suggests that older voters are less likely to favour Labour. To appreciate the scale of these differences, the untransformed values suggest that 45 per cent of the stereotypical voters favour Labour, while this increases to 92 per cent for unemployed 20-year olds living in property rented from the local authority, and declines to 14 per cent for 60-year old members of the private sector salariat.[43]

Thus far these findings accord well with conventional individual-level analyses of the British electorate. We have found clear evidence of: a class cleavage, with the working class more pro-Labour than the middle class; a consumption and a production sector cleavage, with those in the public sector more pro-Labour than those in the private sector; and an age effect, with older voters more Conservative than younger voters.[44] However, we find no evidence here to support the views of those who argue that apparent 'place effects' are only the product of many individual effects.[45] For when we turn to the random

[43] The estimates are obtained by inserting the relevant data into the model. Thus, to take a complex illustration based on Model B, one estimated logit is: -0.20 [the constant term] + 0.84 [if the respondent is unemployed] + 1.30 [if the respondent is a council tenant] + $(-0.02 * -26)$ [the differential for a 20-year-old] + 1.43 [the differential for living in Blaenau Gwent]. This sums to +3.89 for an unemployed, 20-year-old council tenant living in Blaenau Gwent. The transformed percentage is obtained as $[\exp(L_{ij})]/[1.0+\exp(L_{ij})]$, which is $[\exp^{3.89}]/[1.0+\exp^{3.89}]$, which is 99 per cent.

[44] On the latter point, see also A. T. Russell, R. J. Johnston and C. J. Pattie, 'Thatcher's Children: Exploring the Links between Age and Political Attitudes', *Political Studies* (forthcoming).

[45] Notably Rose and McAllister, *The Loyalties of Voters*.

terms, although the level-2 variance has declined substantially with the inclusion of level-1 fixed terms, it is still significantly different from zero: places differ in their voting patterns even when allowance is made for the characteristics of the voters living in these places. Thus the findings presented here further enhance the case that similar people living in different types of place do tend to vote differently.[46]

The twenty largest positive and negative logits are given in Table 4, and the general reduction in deviations from the average in comparison to Model A is clear. These values imply that, for stereotypical voters, the extremes vary around the average of 45 per cent in favour of Labour in Dulwich,[47] from 77 per cent in Blaenau Gwent to 26 per cent in Spelthorne.[48]

Model B is a random intercepts model and as such the constituency random terms represent overall differentials for each type of voter. Thus, it can be estimated that the percentage of unemployed 20-year olds living in rented accommodation who favour Labour varies at the extremes from 99 in Blaenau Gwent to 83 in Spelthorne. For 60-year old members of the private sector salariat, the percentage for Labour varies around the average of 14, from 7 in Spelthorne to 41 in Blaenau Gwent. In both cases we may be predicting (well) beyond the range of the sample data! Moreover, with only a small number of voters in a constituency there is a need to take sampling fluctuations into account when making these comparisons. This is achieved by calculating the ratio of the estimate to its comparative standard error.[49] The only constituencies for Model B in Table 4 that have a ratio in excess of 2.0 are associated with the eleven highest positive logits. Even so, the implications are clear: people with the same individual characteristics seem to be voting very differently in different constituencies. There is a clear constituency effect.

The nature of that effect can be specified more clearly, however, by further modelling which relaxes the constraints on fixed slopes of Model B and which introduces further variables at level 2. *Models C and D* go some way towards this: they include random slope terms as the effects for age are allowed to vary at level-2. The only difference in the specification of the two models is that the level-1 variance is constrained to 1 in Model C and is estimated from

[46] See, for example, R. J. Johnston, 'The Geography of the Working Class and the Geography of the Labour Vote in England, 1983', *Political Geography Quarterly*, 6 (1987), 7–16; I. McAllister, 'Social Context, Turnout and the Vote', *Political Geography Quarterly*, 6 (1987), 17–30; I. McAllister 'Comment on Johnston', *Political Geography Quarterly*, 6 (1987), 45–50; R. J. Johnston and C. J. Pattie, 'Family Background, Ascribed Characteristics, Political Attitudes and Regional Variations in Voting within England, 1983', *Political Geography Quarterly*, 6 (1987), 347–9.

[47] Dulwich has a zero logit. This does not mean that it is an extremely marginal seat but that its voters do not differ from the countrywide average pattern in their relative support for Conservative and Labour.

[48] These percentages are computed as illustrated in fn. 42.

[49] Jones, 'Specifying and Estimating'.

TABLE 5 *Twenty Largest Constituency Differences: Model D**

Constituency	Intercepts	Constituency	Age-slopes
Cambridgeshire South West	−1.01	Wallsend	−0.020
Spelthorne	−0.98	Paisley South	−0.018
Peterborough	−0.95	Monmouth	−0.017
Taunton	−0.90	Erewash	−0.017
Hexham	−0.89	Northampton South	−0.014
Walsall North	−0.86	East Kilbride	−0.014
Beverley	−0.86	Banbury	−0.014
Newbury	−0.85	Kingswood	−0.014
Woking	−0.83	Birmingham Sparkbrook	−0.014
Gloucestershire West	−0.82	Tooting	−0.014
Thurrock	−0.82	Stalybridge and Hyde	−0.013
Cheltenham	−0.80	Bolton West	−0.013
Tayside North	−0.80	Liverpool West Derby	−0.012
Daventry	−0.78	Wallasey	−0.012
Ravensbourne	−0.78	Sherwood	−0.012
Skipton and Ripon	−0.77	Hendon South	−0.012
Bexleyheath	−0.76	Lewisham Deptford	−0.012
Bosworth	−0.76	Sheffield Brightside	−0.011
Dorset South	−0.76	Pendle	−0.011
Saffron Walden	−0.75	Newham North West	−0.011
Livingston	0.86	Poole	0.011
Hull East	0.86	Guildford	0.011
Newham North West	0.88	Skipton and Ripon	0.011
Sheffield Brightside	0.97	Norwich North	0.012
Durham North	0.98	Chingford	0.012
Birkenhead	0.99	Dudley West	0.013
Liverpool Walton	1.02	Leeds South and Morley	0.013
Keighley	1.06	Hertfordshire North	0.013
Merthyr Tydfil and Rhymney	1.06	Horsham	0.013
St Helens North	1.08	Leeds East	0.014
Leigh	1.17	Newark	0.014
Preston	1.20	Shipley	0.014
Bishop Auckland	1.21	Havant	0.015
Monklands East	1.21	Windsor and Maidenhead	0.015
Huddersfield	1.24	Gedling	0.015
Gower	1.25	Loughborough	0.016
Sheffield Attercliffe	1.38	Leeds North West	0.018
Ogmore	1.38	Norfolk Mid	0.019
Renfrew West and Inverclyde	1.43	Luton North	0.020
Blaenau Gwent	1.65	Luton South	0.023

**Model D*: model in which the relationship between voting Labour and age is allowed to vary between constituencies.

the data in Model D. In fact, the level-1 variance estimate is significantly lower than 1, suggesting that we are dealing with ‘under-dispersion’, which is thought to be caused by further clustering that has not been accommodated in the model.[50] Comparing the outcomes of models B, C and D, there is very little change in the fixed estimates: the interpretation is unaltered. In the random part, the variance for random slopes for age is not significant in C (with a Z ratio of 0.1) but approaches significance in D ($Z = 1.9$), when the estimate increases fifteen-fold. Table 5 lists the twenty largest positive and negative effects for both the differential intercepts and slopes for Model D. The decrease in the log-odds with age is steepest in Wallsend, suggesting that older voters are even more Conservative there than elsewhere. Furthermore, the size of the random terms, in comparison to the fixed slope, suggests that there are five places with a positive relationship between age and voting Labour (Loughborough, Leeds North West, Mid Norfolk, Luton North, Luton South). In these five constituencies, it would seem, the relationship between age and voting is the inverse of the national trend, with older voters more likely to support Labour, other things being equal, than younger voters. With such low levels of significance there is an extreme danger of over-interpretation. However, the similarity of the results for the two Luton constituencies suggest that there may be something worth pursuing with a larger sample.[51]

A range of other models has been tried.[52] Table 6 provides results for a series of three-level models (voters in constituencies in regions). The estimates of the fixed parts are very similar to the two-level models and need no further comment. However, the three-level random intercept *Model E* suggests that there are significant differences between regions, and between constituencies within regions, even when voter characteristics are taken into account. All twenty-two regional random intercepts are given in Table 7, together with the ratio of these estimates to their comparative standard error. The most pro-Labour region is clearly Industrial South Wales, while the most anti- is the Outer Metropolitan region, both differences being significant. Broadly, the

[50] 'Under-dispersion' occurs when the level-1 random term is significantly less than 1.00. There is thus less variation at the voter-level than would be expected if the conditional probability of voting Labour followed a pure binomial distribution. See Goldstein, 'Nonlinear Multilevel Models'.

[51] This may also be because all of the working-class respondents in those constituencies (more exactly, the polling districts selected within them) are relatively old and younger residents are not working-class. The result may be an artefact of the data collection, however, specifically an interviewer effect. One study found that, using an ML approach, not only did some interviewers find more disability than others, but that they also elicited different relationships with age. See R.O. Wiggins, N. Longford and C. O'Muircheartaigh, *A Variance Components Approach to Interviewer Effects* (London: Joint Centre for Survey Methods Research, Working Paper 2, 1990).

[52] Software crashes were experienced when the class variables and the tenure categories were allowed to be random at the constituency level. This is caused by insufficient data to estimate the covariances in the random part, as this requires that each type of class or category be present in each constituency; with only around nine respondents in each seat, this was not always the case. It is possible to estimate the level-2 variances while constraining the covariances to zero, but this is not a recommended procedure when the variables that are being allowed to vary at the higher level form a dummy variable set: see Longford, *Manual for VARCL.* One other model, in which the effects for unemployment were allowed to vary between constituencies, did successfully converge, but the associated variance did not approach significance.

TABLE 6 *Summary of Estimated Three-Level Models: Individuals in Constituencies in Regions**

	Model E		Model F		Model G	
	Estim.	Z	Estim.	Z	Estim.	Z
FIXED						
CON	−0.08	–	−0.08	–	−0.07	–
UNEMP	0.75	3.6	0.76	3.6	0.76	3.7
AGE	−0.02	−6.5	−0.02	−5.4	−0.02	−6.5
PRISAL	−1.29	−9.0	−1.29	−8.9	−1.3	−9.1
NONMAN	−0.81	−7.3	−0.80	−7.2	−0.81	−7.3
BOURG	−1.20	−6.8	−1.19	−6.8	−1.19	−6.8
FORSUP	−0.30	−1.6	−0.31	−1.7	−0.29	−1.6
PUBSAL	−0.47	−3.1	−0.47	−3.1	−0.47	−3.1
COUNC	1.23	10.5	1.24	10.6	1.22	10.4
PRITEN	0.55	3.5	0.54	3.4	0.48	2.2
RANDOM						
Level 3						
CON	0.35	2.8	0.36	2.8	0.36	2.8
AGE/CON			−0.002	−1.0		
AGE			0.00008	1.2		
COUNC/CON					0	0
COUNC					0	0
PRITEN/CON					−0.13	−0.9
PRITEN/COUNC					0	0
PRITEN					0.44	1.5
Level 2						
CON	0.23	3.6	0.22	3.6	0.23	3.6
Level 1						
CON	0.79	–	0.79	–	0.78	–

* *Model E*: three-level model with unconstrained level-1 variance and nine individual-level predictors.
Model F: as E, but with complex random part in which the age relationship is allowed to vary between constituencies.
Model G: as E, but with complex random part in which relationships with tenure are allowed to vary between constituencies.
All estimates on a logit scale.

results of the model confirm the north–south, urban–rural divide of the British electorate identified by several commentators. Labour achieved significantly better results than its national average in the large industrial conurbations of Scotland and the north of England, and in South Wales, while the Conservatives did best in the central and southern English regions of the East Midlands, the Outer Metropolitan area, and in rural East Anglia and the South West of England. Although some other regions approached significance (notably West Yorkshire as a potentially pro-Labour region, and Outer London as a Conservative stronghold) the remainder of the twenty-two regions did not differ significantly from the national average.

TABLE 7 *Random Intercepts and Ratios of Estimates to Comparative Standard Errors, Model E**

Economic Region	Intercepts	Z Ratios
Strathclyde	0.78	3.3
East-Central Scotland	0.04	0.2
Rural Scotland	−0.46	−1.3
Rural North	−0.35	−1.7
Industrial North-East	0.31	1.4
Merseyside	0.52	2.0
Greater Manchester	0.13	0.5
Rest of North-West	0.23	1.1
West Yorkshire	0.48	1.8
South Yorkshire	0.88	3.3
Rural Wales	0.10	0.3
Industrial South Wales	1.23	5.2
West Midland Conurbation	−0.10	−0.5
Rest of West Midlands	−0.05	−0.2
East Midlands	−0.47	−2.5
East Anglia	−0.74	−3.0
Devon and Cornwall	−0.60	−2.2
Wessex	−0.69	−3.4
Inner London	0.25	1.0
Outer London	−0.33	−1.8
Outer Metropolitan	−0.82	−5.5
Outer South-East	−0.33	−1.8

* *Model E*: differential regional logits after allowing for nine individual-level predictors.

To appreciate the scale of these inter-regional differences, Table 8 provides predicted Labour percentages of the two-party vote for the three types of voter in each region. In Industrial South Wales, for instance, Model E predicts that 76 per cent of our stereotypical 46-year old, working-class, employed voters who own their own homes would have voted Labour, given a straight choice between that party and the Conservatives. That percentage rises to 94 of 20-year-old unemployed South Wales council tenants, but falls to only 41 per cent among the region's 60-year-old private-sector salariat voters. By contrast, in the Outer Metropolitan region, given the same choice between Labour and the Conservatives, the model predicts that only 29 per cent of the stereotypical voters would support Labour, falling to a mere 8 per cent of the 60-year old private salariat, while 69 per cent of unemployed 20-year olds would do so.

A number of three-level random slope models were also fitted to the data but it must be remembered that this requires the estimation of parameters on the basis of only twenty-two observations. In *Model F* the relationship with age is allowed to vary between regions, but is found not to be significant. Figure 3 portrays the predicted relationships, and there is a general impression of a random intercepts model: by and large, the slopes are roughly parallel, suggesting that being in a particular region has the same effect on all voters,

TABLE 8 *Estimated Percentage Who Vote Labour, Model E*

Economic Region	Stereotypical voter* I	II	III
Strathclyde	67	91	31
East-Central Scotland	49	83	17
Rural Scotland	37	75	11
Rural North	39	77	12
Industrial North-East	56	87	22
Merseyside	61	89	25
Greater Manchester	51	85	19
Rest of North-West	54	86	20
West Yorkshire	60	89	25
South Yorkshire	69	92	33
Rural Wales	51	84	18
Industrial South Wales	76	94	41
West Midland Conurbation	45	81	15
Rest of West Midlands	47	82	16
East Midlands	37	75	11
East Anglia	31	70	9
Devon and Cornwall	34	72	10
Wessex	32	71	9
Inner London	54	86	21
Outer London	40	78	13
Outer Metropolitan	29	68	8
Outer South-East	40	78	13

*I: 46-year old, employed, working-class owner-occupier.
II: 20-year old, unemployed, council-house tenant.
III: 60-year old, private-sector salariat.

irrespective of their age. In *Model G*, both differential tenure categories are allowed to vary between regions. The variance of the differential effects for the local authority tenure are estimated at zero and although there is substantial level-3 variation associated with the 'other' tenure category, this does not reach significance. Table 9 lists the predictions of Labour voting from this model by our stereotypical voters in each of the twenty-two regions, for each of the three tenure categories; these are presented graphically in Figure 4. There are some interesting differences: Strathclyde is the most pro-Labour region for the 'other tenure' category, for example, whereas the Rural North has the lowest percentage favouring Labour. In the Outer Metropolitan Area, on the other hand, there are very low Labour percentages in the owner-occupier and local authority tenure categories but the region has the fourth highest percentage of the 'other tenure' category favouring Labour. Thus there are major regional variations in the propensity to vote Labour among occupants of the three separate tenure categories, again indicative of place effects that are independent of the characteristics of the individual voters.

A number of models was also fitted which allow the class cleavage variables

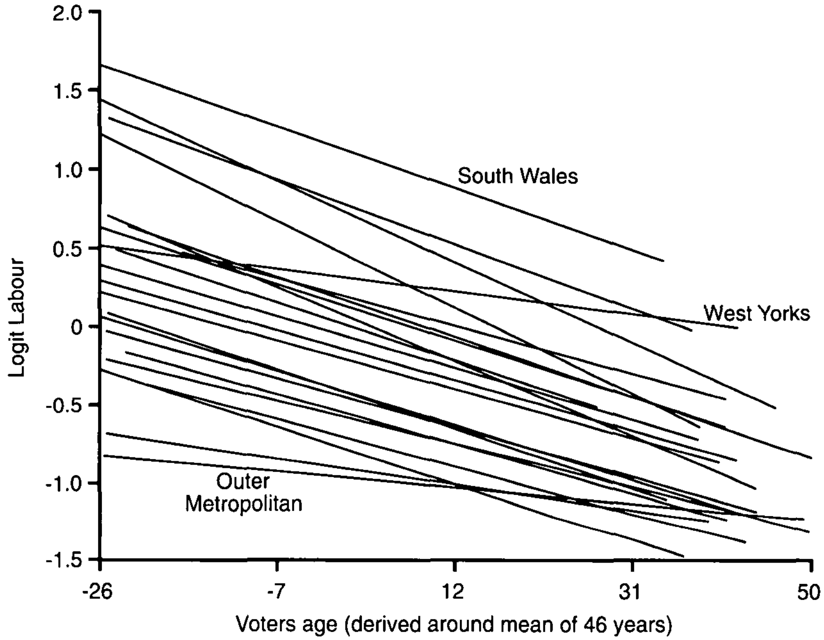


Fig. 3. The relationship between age and the logit for Labour voting in a level-3 model (Model F in the text), showing the variation in the relationship between the twenty-two regions

to vary at regional level; various combinations of the variables were tried but none even approaches significance. This implies that the same regional logit differences apply for each class. Consequently, Table 10 shows the predictions of Labour voting by class for each region on the basis of the random intercepts Model E, for 46-year old employed owner-occupiers. Here we see clear evidence that the class cleavage varies from region to region. At the extremes, Labour did better than the national average among all classes in Industrial South Wales, and worse than the average in the Outer Metropolitan region, even when the individual characteristics of voters and the contexts of particular constituencies were taken into account. Thus, for example, 76 per cent of the working-class 46-year-old employed owner-occupiers in Industrial South Wales were predicted as voting Labour: even among the most anti-Labour classes, the petty bourgeoisie and the private-sector salariat, almost half (49 and 46 per cent respectively) of this group were predicted as Labour supporters. At the other extreme, however, in the Outer Metropolitan region, Model E predicts that only 29 per cent of working-class 46-year-old employed owner-occupiers would vote Labour, falling to only 10 and 11 per cent respectively for the private-sector salariat and the petty bourgeoisie. As would be expected from a random intercepts model, the overall pattern of the class cleavage is constant: the working

TABLE 9 *Estimated Percentage According to Tenure of 46-Year Old, Employed, Working-Class Who Vote Labour, Model G*

Economic region	Owner occupier	Other	Local Authority
Strathclyde	65	81	86
East-Central Scotland	49	66	76
Rural Scotland	38	48	67
Rural North	43	32*	72
Industrial North-East	57	56	82
Merseyside	61	67	84
Greater Manchester	51	66	78
Rest of North-West	54	60	80
West Yorkshire	60	71	83
South Yorkshire	69	75	88
Rural Wales	51	60	78
Industrial South Wales	76	78	92
West Midland Conurbation	46	56	74
Rest of West Midlands	48	55	75
East Midlands	37	49	66
East Anglia	31	42	60
Devon and Cornwall	34	46	63
Wessex	32	47	61
Inner London	54	66	80
Outer London	39	62	68
Outer Metropolitan	27	73*	55
Outer South-East	40	52	70

* Regions in which the differential estimates for the 'other' category are over twice their comparative standard errors.

class is more pro-Labour than the middle classes, for instance. But the dimensions of that cleavage vary substantially between regions, further evidence that 'place matters'.[53]

All of the models discussed above, with the exception of Model A, have used the same level-1 variables in the fixed part. The final set of models to be discussed here includes higher level variables in the fixed part to account for the differences between places that remain after taking account of individual voter characteristics. Three variables are used: unemployment in 1983; change in unemployment 1983–86; and percentage employed in mining, 1981. The first two represent a major element of the uneven development (or economic restructuring) that characterized the British economy in the early 1980s and was thought responsible for the expansion of the north–south divide;[54] the third

[53] R. J. Johnston, *A Question of Place* (Oxford: Basil Blackwell, 1991).

[54] R. J. Johnston and C. J. Pattie, 'Class Dealignment and the Regional Polarisation of Voting Patterns in Great Britain, 1964–1987', *Political Geography*, 11 (1992), 73–86; R. J. Johnston, C. J. Pattie and A. T. Russell, 'Dealignment, Spatial Polarisation and Economic Voting', *European Journal of Political Research*, 20 (1992), forthcoming.

TABLE 10 *Estimated Percentage According to Occupational Class of 46-Year Old, Employed, Owner-Occupiers Who Vote Labour: Model E*

Economic region	Working class	Private sector salariat	Routine non-manual	Petty bourgeoise	Foremen/ Supervisors	Public sector salariat
Strathclyde	67	36	47	38	60	56
East-Central Scotland	49	21	30	23	42	38
Rural Scotland	37	14	21	15	30	27
Rural North	39	15	22	16	33	29
Industrial North-East	56	26	36	27	48	44
Merseyside	61	30	41	32	54	49
Greater Manchester	51	22	32	24	44	40
Rest of North-West	54	24	34	26	46	42
West Yorkshire	60	29	40	31	53	48
South Yorkshire	69	38	50	40	62	58
Rural Wales	51	22	31	24	43	39
Industrial South Wales	76	46	58	49	70	66
West Midland Conurbation	45	19	27	20	38	34
Rest of West Midlands	47	19	28	21	40	35
East Midlands	37	14	21	15	30	27
East Anglia	31	11	17	12	25	22
Devon and Cornwall	34	12	18	13	27	24
Wessex	32	11	17	12	26	22
Inner London	54	25	35	26	47	42
Outer London	40	15	23	17	33	29
Outer Metropolitan	29	10	15	11	23	20
Outer South-East	40	15	23	17	33	29

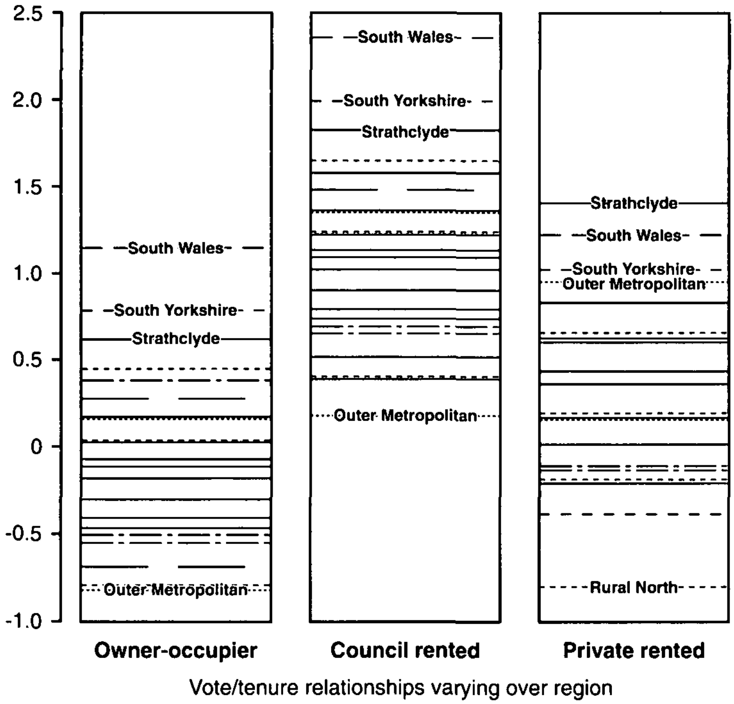


Fig. 4. The inter-regional differences in the propensity to vote Labour among the three tenure categories and the twenty-two regions, according to a three-level model (G)

represents the major areas of Labour support, which were affected by the differential response to the 1984–85 NUM strike.[55] All three variables are percentages which have been deviated about the mean for all constituencies. Three exploratory models were fitted using two of these constituency variables, one using two levels and the other two using all three levels: the results are in Table 11.

Model H has the same form as Model B except that the level-1 random term is unconstrained and the three constituency variables are included in the fixed part. In comparison with the other two-level models fitted (Table 3) the level-1 fixed terms are essentially unchanged. However, all three level-2 variables are significant, suggesting that individuals who lived in areas with high

[55] M. J. Griffiths and R. J. Johnston, 'What's in a Place? An Approach to the Concept of Place as Illustrated by the British National Union of Mineworkers' Strike, 1984–85', *Antipode*, 23 (1991), 185–213.

TABLE 11 *Summary of Estimated Models with Constituency Variables in the Fixed Part**

	Model H		Model I		Model J	
	Estim.	Z	Estim.	Z	Estim.	Z
FIXED						
Level 1						
CON	−0.21	–	−0.13	–	−0.13	–
UNEMP	0.75	3.2	0.73	3.2	0.72	3.1
AGE	−0.02	−6.0	−0.02	−6.2	−0.02	−6.1
PRISAL	−1.34	−8.1	−1.35	−8.3	−1.35	−8.2
NONMAN	−0.83	−6.7	−0.85	−6.9	−0.85	−6.8
BOURG	−1.29	−6.4	−1.31	−6.7	−1.30	−6.5
FORSUP	−0.37	−1.8	−0.36	−1.8	−0.36	−1.7
PUBSAL	−0.53	−3.1	−0.53	−3.1	−0.51	−3.0
COUNC	1.23	9.5	1.26	9.9	1.25	9.7
PRITEN	0.55	3.1	0.55	3.2	0.56	3.2
Level 2						
ConEmp	0.07	4.8	0.06	4.2	0.07	4.6
ConCha	0.03	3.5	0.02	2.6	0.02	2.9
ConMine	0.04	2.0	0.03	1.5	0.06	1.6
RANDOM						
Level 3						
CON			0.13	2.2	0.08	1.8
CON/ConMine					0.01	0.4
ConMine					0.01	1.8
Level 2						
CON	0.34	4.2	0.21	3.0	0.16	2.4
Level 1						
CON	0.90	–	0.88	–		

**Model H*: two-level model as B, but includes three constituency variables at level 2 and unconstrained level-1 variance.
Model I: three-level model as E, but includes three constituency variables at level 2.
Model J: as I, but allows the constituency relationship with mining to vary between regions.

unemployment, in areas that experienced (relatively) worsening unemployment in the mid-1980s, and in mining constituencies were more likely to vote Labour, irrespective of their individual characteristics. This is reflected in the decline in the level-2 variance. Table 12 (which should be compared with Table 4 on p. 362) shows the constituency differentials.

Model I is a three-level random intercepts model of the same form as Model E with the addition of the three constituency variables. Significant fixed level-2 estimates for unemployment and change in unemployment are found and consequently there is a reduction in both the high-level random terms, although this is most marked for the regional-level variance. The differential regional preferences for Labour can thus in part be accounted for statistically by how

TABLE 12 *Twenty Largest Differences for Constituencies, Model H**

Constituency	Logit
Bassetlaw	−0.92
Newcastle on Tyne Central	−0.87
Tynemouth	−0.66
Walsall North	−0.65
Halton	−0.64
Peterborough	−0.62
Chesterfield	−0.60
Wolverhampton South West	−0.55
Midlothian	−0.55
Beverley	−0.54
Nottingham North	−0.54
Thurrock	−0.52
Hexham	−0.51
Taunton	−0.50
Truro	−0.49
Tayside North	−0.49
Cheltenham	−0.49
Staffordshire South	−0.48
Devon West and Torridge	−0.47
Wallasey	−0.46
Swindon	0.52
Sheffield Hillsborough	0.52
Dover	0.55
Livingston	0.55
Durham North	0.56
Staffordshire Moorlands	0.58
St Helens North	0.62
Preston	0.62
Monklands East	0.66
Leigh	0.66
Pendle	0.66
Blaenau Gwent	0.78
Wealden	0.80
Sheffield Attercliffe	0.83
Ogmore	0.83
Huddersfield	0.84
Gower	0.84
Keighley	0.89
Bishop Auckland	0.90
Renfrew West and Inverclyde	1.04

**Model H*: differential logits for two-level model that includes nine individual-level predictors at level 1 and three constituency-level predictors at level 2.

the constituencies in the regions are faring economically. Table 13 provides the regional differentials for Models I and E: interestingly, the Industrial North

TABLE 13 *Random Intercepts for Models I and E**

Economic Region	Model H	Model E
Strathclyde	0.31	0.78
East-Central Scotland	−0.09	0.04
Rural Scotland	−0.25	−0.46
Rural North	−0.14	−0.35
Industrial North-East	−0.10	0.31
Merseyside	0.04	0.52
Greater Manchester	0.07	0.13
Rest of North-West	0.18	0.23
West Yorkshire	0.38	0.48
South Yorkshire	0.37	0.88
Rural Wales	0.10	0.10
Industrial South Wales	0.80	1.23
West Midland Conurbation	−0.21	−0.10
Rest of West Midlands	0.15	−0.05
East Midlands	−0.45	−0.47
East Anglia	−0.27	−0.74
Devon and Cornwall	−0.47	−0.60
Wessex	−0.31	−0.69
Inner London	−0.01	0.25
Outer London	−0.04	−0.33
Outer Metropolitan	−0.23	−0.82
Outer South-East	0.15	−0.33

**Model I*: regional differential logits for three-level model that includes individual- and constituency-level predictors.
Model E: regional differential logits for three-level model that includes only individual-level predictors.

East is shown as relatively anti-Labour and the Outer South East as pro-Labour when economic prosperity is taken into account.[56]

Finally, in *Model J* the constituency-level relationship between mining and voting Labour is allowed to be random at the regional level. The fixed terms are very similar to those in Model I, but remembering that the three level-3 random parameters are estimated on only twenty-two regions, there does appear to be something to explore further. Figure 5 shows the relationships for each region between the logit of voting Labour (after allowing for constituencies' composition in terms of age, class and tenure) and the percentage of miners in each constituency's work-force. The overall expected positive fixed slope conceals interesting inter-regional differences. The most marked is between

[56] This may well indicate that there is a greater 'bed-rock' support for Labour in the South East than in the Industrial North East. Once individual social background, constituency milieux influences and local economic voting have been screened out we have accounted for all of the Labour voting in the Industrial North East, whereas the Outer South East still has some pro-Labour voters not accounted for by those influences. This might possibly be the outcome of embourgeoisement: see R. J. Johnston, '"Embourgeoisement", the Property-Owning Democracy and Ecological Models of Voting in England', *British Journal of Political Science*, 11 (1981), 499–503).

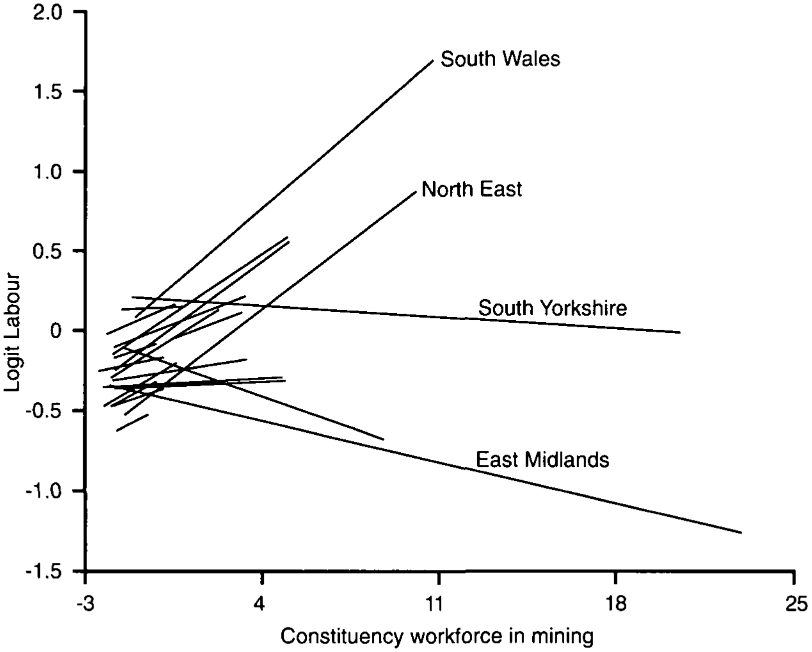


Fig. 5. The variation in the relationship between the percentage of miners in a constituency and the propensity of individuals in the constituency to vote Labour, between the twenty-two regions according to a three-level model (J)

South Wales and the East Midlands: the former is a pro-Labour area and support for that party increases the greater the importance of mining in a local constituency's economy; the East Midlands contains some of the most anti-Labour constituencies of the mining regions, however, reflecting the local culture there, as clearly demonstrated during the 1984–85 strike.[57] Both South Wales and the Industrial North East have positive slopes, though with a lower base for the latter region: the more miners in a constituency the greater the Labour vote, *ceteris paribus*. In South Yorkshire, however, the support for Labour did not vary with the mining percentage, and in the East Midlands the larger the mining component of the local economy the smaller the Labour vote!

CONCLUSIONS

The analyses laid out above clearly demonstrate the value of a multi-level modelling approach for studies of electoral behaviour. In particular, much of the debate on the nature of contextual effects has been dogged by the interpretation of regression models. For some, the crucial influences on voters' choices are their individual attributes: apparent contextual effects, they argue, are only

[57] Johnston, *A Question of Place*; Griffiths and Johnston, 'What's in a Place?'

the result of the aggregation of many individuals and their individual effects (level-1 effects, in our discussion above) in particular places. For others, however, this fundamentally misses the point, as many apparently 'individual' attributes are learned in particular contexts. Constituency and regional effects (level-2 and level-3 effects above) are not 'add-on' elements to be included after individual effects are taken into account: rather, they are implicated in the 'individual' effects from the outset.

It has been difficult, if not impossible, to resolve this argument using conventional regression methods, since least squares regression cannot deal easily with highly autocorrelated data and cannot 'partition' variance in data into hierarchical levels. Multi-level methods, however, are designed to do just that. In this initial article, therefore, we have been able to demonstrate, more rigorously than before, that 'individual effects' are not the sole determinants of voting behaviour. Voting is done by individuals, but they do so in particular contexts, and those contexts influence voters in their own right.

The Potential of Multi-Level Modelling as General Framework

Multi-level modelling provides a rich set of models for the empirical analysis of voting behaviour. The present exposition and illustration has dealt with cross-sectional data with binary response only. Table 14 gives some indication of the diverse problems that are appropriately tackled within the general framework of the multi-level approach,[58] and suggests how the utility of the approach might be expanded.

TABLE 14 *Research Problems in a Multi-Level Framework*

Problem	Approach	Levels
Polytomous choice	Multinomial logit	1 choices 2 individuals 3 constituencies
Stability of contextual effects over time	Include polynomials of time	1 individuals 2 time 3 constituencies
Switching behaviour of individuals	Repeated measures design	1 time 2 individuals 3 constituencies
Alternative 'groupings' at higher level	Cross-classified models	1 individuals 2a constituency 2b labour-market area

[58] Jones, *Multilevel Models for Geographical Research.*

The first problem concerns our treatment of voting as a binary choice. Goldstein has developed the estimation theory for the multinomial logit which can simultaneously model responses when there are more than two possible outcomes.[59] This is achieved by specifying level 1 as the different choice sets, level 2 as the individual, and level 3 as the constituency. ML3 macros are currently being prototyped for this problem in an exceedingly general way such that there can be sets of discrete choices for the response variables as well as continuous variables in a multivariate model. Care is needed in the specification of such models, however, as it is an underlying assumption of the multinomial that the choice set is the same for each voter. Obviously, this is not the case, with the nationalist parties, for example, only contesting seats in Scotland and Wales.

The second research area concerns the stability of the contextual effects over time. The aim is to model a number of cross-sectional data sets collected at each election to see how consistent are the place effects through time. Such research has been undertaken using multi-level models in educational research, with Willms and Raudenbush distinguishing 'genuine' variation in higher-level, school effects, from 'sampling' fluctuations.[60] Very usefully, not all constituencies need to be represented at each time point. The third problem is to model the changes in individual voting behaviour over time. This sort of repeated-measurement longitudinal design is specified in a multi-level model by taking level 1 to be the measurement occasion, level 2 to be the individual, and level 3 as the constituency. The ability of the multi-level estimates to borrow strength is again useful in relaxing the restrictions on the number and spacing of these measurements.

The final problem concerns the differing levels at which the contextual effects may be operating. It is unnecessarily restrictive for the higher-level structures to be hierarchical and nest within each other. For example, it may be suggested that at level 2, the contextual influences on voting behaviour depend not only on where the voters live and vote (the constituency) but also on where they work, which may be represented by functional, labour-market areas. The appropriate approach is then cross-classified multi-level models for which Goldstein has developed a valid estimation strategy which is currently being implemented in his ML3 software.[61]

The analysis of complex survey designs is currently a very active area of research.[62] Multi-level models which take into account the inherent structure of the data overcome a number of technical difficulties. Moreover, by working simultaneously at the individual and contextual levels, they provide a means

[59] Personal communication with Prof. Goldstein.

[60] J. D. Willms and S. W. Raudenbusch, 'A Longitudinal Hierarchical Linear Model for Estimating School Effects and their Stability', *Journal of Educational Measurement*, 26 (1989), 209–32.

[61] Goldstein, *Multilevel Models in Educational and Social Research*.

[62] C. J. Skinner, D. Holt and T. M. F. Smith, eds, *Analysis of Data from Complex Surveys* (Chichester: John Wiley, 1989).

of 'thick' quantitative description. The idiosyncratic characteristics of the individual voter and her 'interaction' with the context in which she votes are retained in a model which still has a capacity for generalization. The complexity of the real world of voters and places is not ignored in the pursuit of a single universal equation.