J. R. Statist. Soc. A (2000)
163, *Part* 1, *pp.* 49–62

Multilevel models for repeated binary outcomes: attitudes and voting over the electoral cycle

Min Yang and Harvey Goldstein

Institute of Education, London, UK

and Anthony Heath

Nuffield College, Oxford, UK

[Received November 1996. Final revision April 1999]

Summary. Models for fitting longitudinal binary responses are explored by using a panel study of voting intentions. A standard multilevel repeated measures logistic model is shown to be inadequate owing to a substantial proportion of respondents who maintain a constant response over time. A multivariate binary response model is shown to be a better fit to the data.

Keywords: Longitudinal binary data; Multilevel; Multivariate multilevel model; Political attitudes; Voting

1. Introduction

The electoral cycle has become an established feature of voting behaviour, both in Britain and in other European countries. After an initial 'honeymoon' between the new Government and the electorate, disillusion often sets in and Government popularity — whether measured by opinion polls, by-elections or mid-term elections such as the European and local elections — tends to decline. In most cases, there is then some recovery in the Government's standing in the run-up to the next general election (Miller *et al.*, 1986; Miller and Mackie, 1973; Reif, 1984; Stray and Silver, 1983). During the 1987–1992 British Parliament, for example, the Conservative Government lost seven by-elections but subsequently won all back at the 1992 general election. Although the Conservatives were much less successful in 1997 than they had been in 1992, their result in the 1997 general election marked a recovery from their lowest point in the electoral cycle at the time of the 1995 local elections.

There are various possible explanations for this pattern. One is that voters make their mid-term decisions on rather different criteria from those that they use at a general election. Thus, in the mid-term, votes at a by-election or at the European election are unlikely to lead to a change in the Government. Voters may on these occasions communicate their dissatisfaction rather than wish to change the Government. This point may hold with even more force for mid-term opinion polls. As Miller and Mackie (1973) suggest, an explanation for the cyclical pattern of Government popularity in opinion polls may be that

> 'the . . . poll series changes in meaning as time passes. The wording of the question remains unaltered but the political context in which it is asked changes, and the replies of the interviewees are responses to a "question in context"'(Miller and Mackie (1973), pages 265–266).

Address for correspondence: Min Yang, Multilevel Models Project, Department of Mathematical Sciences, Institute of Education, University of London, 20 Bedford Way, London, WC1H 0AL, UK.
E-mail: m.yang@mentor.ioe.ac.uk

 0964–1998/00/163049

Gelman and King (1993) have provided a more detailed theory about the way in which the opinion poll series changes in meaning as time passes. Writing about American Presidential campaigns, they asked why the early opinion polls are such poor predictors of the eventual outcome. They suggested that, at the start of the campaign (which in America is substantially longer than the four weeks of the usual British campaign) voters do not have the information that is necessary to make enlightened voting decisions. Their responses to pollsters early in the campaign are thus based on unenlightened preferences, using whatever information they happen to have to hand about the candidates. Voters then acquire the information that is needed to make enlightened decisions over the course of the campaign, and by polling day can base their decisions on what Gelman and King (1993) termed their 'fundamental variables', i.e. the voters learn how the candidates' policies relate to their own ideologies. Fundamental variables such as the voters' ideologies thus come to acquire greater weight as the campaign progresses.

Although there are important institutional differences between the American Presidential campaigns studied by Gelman and King and the British electoral campaigns, similar processes may be at work here. Thus voters' responses to opinion pollsters in the middle of the electoral cycle may be based primarily on the 'headline' information that they have at hand from the mass media about the candidates and parties. But on polling day, when they must make a more consequential decision about which party should govern, they may be more influenced by their own long-term underlying values and interests. Our hypothesis, then, is that variables such as the voters' ideologies will have relatively greater weight on their actual voting decisions in general elections than they do on decisions in mid-term elections or on voting intentions conveyed to opinion pollsters. The latter, we suspect, will be more influenced by the information which the voters have at hand from the mass media about current political stories and events.

We might also expect that contextual variables at the constituency level will be more important at general elections than during the mid-term. For example, tactical voting depends on the interplay between the perceived political situation in the voter's constituency and the voter's personal preferences for the parties. There is some evidence (and good theoretical reasons) for thinking that tactical voting is more common at general elections than during the middle of the electoral cycle (McLean *et al.*, 1996). In principle, therefore, we would like to be able to incorporate the constituency level within our modelling of the electoral cycle (see Section 6).

There is some evidence (and good theoretical reasons) for thinking that tactical voting is more common at general elections than during the middle of the electoral cycle (McLean *et al.*, 1996) and it is unlikely that tactical considerations will play much part in answers to a question on mid-term voting intention. We therefore expect there to be less constituency variation in the middle of the cycle and accordingly we would like to be able to incorporate the constituency level within our modelling of the electoral cycle (see Section 6).

Gelman and King (1993) used a series of independent random samples conducted at different stages of the campaign to test their hypotheses. A more efficient design, however, for understanding changes in voters' behaviour is to use a panel study, with repeated observations on the same respondents. In the present paper, we use a three-wave panel study covering a complete electoral cycle from 1983 to 1987 to illustrate the modelling procedures. There are three important features of the structure of the data set:

(a) a hierarchical structure with voters nested within constituency and years nested within a voter;

(b) repeated dependent binary outcomes (vote or voting intention);
(c) time-dependent covariates representing voters' ideologies and perceptions of the parties and their leaders.

The dependence problem can be tackled by using the *arbitrary multinomial* (Cox, 1972). A multinomial distribution is fitted to the $2^k - 1$ combinations resulting from the k binary outcomes and in our case $k = 3$ (years). The lowest level of the hierarchy is defined by the three-element multinomial vector resulting from the reformulation. This model cannot easily accommodate time-dependent covariates which are one level lower than the new multinomial response, and as Cox pointed out this model gives little insight into the structure of the data (Cox (1972), page 115). Alternatively the *multivariate logistic model* (Cox, 1972) which includes covariance terms between the outcomes to take up the dependence may be useful, since time-dependent covariates can also be incorporated.

Korn and Whittemore (1979) analysed the data from a panel study of acute health effects of air pollution, using a logistic growth model over time for each individual, then accumulating the estimates and responses over individuals by means of weighting. The intercept term of the logistic model was modelled by a time series expression. The maximum likelihood (ML) estimates were based on the likelihood for all occasions for each individual. The summary of these estimates over all individuals was carried out as the second stage of the analysis. Time-dependent covariates were accommodated. This model worked on a long time series for each subject, and no further multilevel structure was considered.

To take into account the clustering and to model the contextual effects found in the data, Goldstein (1986) proposed the multilevel model using iterative generalized least squares estimation. Under normality assumptions this leads to ML estimates. For repeated responses over occasions the model can be extended naturally by adding a further level at the bottom of the data structure, giving three levels in all. Effects of respondent level and occasion level covariates can be estimated in the fixed part of the model.

To extend this method to the case of a repeated binary response variable, we may use a generalized linear model formulation (Goldstein and Rasbash, 1996; Diggle *et al.*, 1994). At the voter level, we can consider modelling the probability of a positive response as a smooth, for example polynomial, function of time. This can be convenient and efficient when the time series is long. However, the standard assumptions of such a model may not be realistic. For example, suppose that we have repeated measures of voting for a sample of individuals where the binary response is whether a person voted for political party A. For many people their probability of voting for this party will be close to 1 or 0. On the linear scale this implies that such individuals have very large positive or negative values, implying that the underlying distribution is at least trimodal, with a large negative and a large positive mode, so that the standard model which, on the linear scale, assumes normality for the distribution across persons will be misspecified.

Another approach is to use a multilevel multivariate logistic model, a development of Cox's multivariate logistic model. Like any multivariate model, the dependence between the responses can be modelled by the covariance structure at the individual level, in this case the *biserial covariance* (Goldstein, 1995).

In this paper we examine two models: a standard three-level repeated measures logistic model and a multivariate two-level logistic model. We compare the results from these with those obtained by applying separate two-level models to each round of the panel study. For binary responses we use the procedures known as penalized quasi-likelihood (PQL) estimation with a second-order Taylor series approximation (Breslow and Clayton, 1993; Goldstein, 1991; Goldstein and Rasbash, 1996) which has been incorporated into the program MLwiN (Rasbash *et al.*, 1999). This estimation procedure can produce biases when there are small numbers of level 1 units per level 2 unit, but in the present case the actual level 2 variance estimates are not large and do not change much when moving from 'marginal quasi-likelihood' (MQL) to 'penalized' estimation. In this situation any biases are expected to be small (Goldstein and Rasbash, 1996). In the case of the binary repeated measures model, which is anyway unsatisfactory as we have pointed out, the PQL procedure did not converge and MQL was used.

2. The data from the 1983–1986–1987 panel study

In this paper we use the 1983–1986–1987 British Election Panel Study. Respondents were interviewed on three occasions: first in 1983 immediately after the general election, second in the autumn of 1986 and third in 1987 immediately after the general election of that year. The panel study thus covers a complete electoral cycle, with one round of interviews taking place between the two general elections (Heath *et al.*, 1996).

The respondents to the first round of interviews were drawn from the 1983 British Election Survey (BES). The 1983 BES was a clustered random sample with 3955 respondents interviewed in 250 constituencies (for full details see Heath *et al.* (1985), appendix I). For reasons of cost, the panel was based on a subset of these respondents. Respondents in 112 of the original constituencies were selected to provide the panel. Of the 1698 respondents selected in this way 869 (52%) completed all three waves. The two main sources of non-response were the difficulty of locating respondents who had moved between 1983 and 1987 (a total of 206 individuals) and the refusal of located respondents to participate in an interview (240 individuals). There were also 47 respondents to the 1983 survey who were found to have died by the time of the 1987 survey, and there were a further 56 who were ill, incapacitated or in hospital at the time of the interview. (For further details see Heath *et al.* (1991), appendix II.) There were also 18 respondents whose votes were coded as missing. Therefore, among the remaining 1680 respondents, 603 had data on only one occasion, 234 had data on two occasions and 843 had data on all three occasions. This leaves us 3600 valid responses to work with. Furthermore, there were 243 (9.5%) votes out of 67 respondents who had missing codes on some explanatory variables. These have been removed from the data set so the numbers of responses used are 1502, 1008 and 846 respectively in the three years (Table 1). Among the voters with complete records, the odds voting for Conservative *versus* others both in 1983 and in 1986 were 216:362, whereas among the voters with only two votes in 1983 and 1986 for Conservative the odds were 57:119, which is an odds ratio of 1.25.

Our response variable is vote (or, in 1986, voting intention). In all three rounds of the panel study, data were collected in standard form on the respondents' political attitudes towards basic issues, their evaluations of the party leaders and their images of the parties. In the 1983 and 1987 rounds of interviews respondents reported how they had voted at the relevant general election, and in the 1986 round respondents were asked what their voting intention was. The detailed questionnaires and scale cards can be found in Heath *et al.* (1991), pages 251–309. Voting intention is conceptually different from reported vote, but this is precisely what we wish to investigate. To simplify the treatment we shall in our analysis dichotomize the response, contrasting Conservative votes with votes for all other parties. Since the substantive theories focus on disillusion with the incumbent Government during the middle of the electoral cycle, this contrast between the incumbent Government and the opposition parties is appropriate.

Table 1. Frequency of respondents by their voting occasions for the Conservative Party in the panel†

3 responses		*2 responses*		*1 response*	
1983–1986–1987	*Frequency*	*1983–1986–1987*	*Frequency*	*1983–1986–1987*	*Frequency*
1–1–1	202 (0.277)	1–1–×	57 (0.199)	1–×–×	221 (0.370)
1–0–0	42 (0.058)	1–×–1	11 (0.038)	×–1–×	5 (0.008)
1–1–0	14 (0.019)	×–1–1	12 (0.042)	×–×–1	2 (0.003)
0–1–0	16 (0.022)	0–0–×	119 (0.416)	0–×–×	340 (0.569)
0–0–1	37 (0.051)	0–×–0	25 (0.087)	×–0–×	24 (0.040)
1–0–1	64 (0.088)	×–0–0	62 (0.217)	×–×–0	6 (0.008)
0–1–1	29 (0.040)				
0–0–0	325 (0.446)				
Total	729	Total	286	Total	598

†1, yes; 0, others; ×, not available; proportions are given in parentheses.

Table 2. Description of explanatory variables

Variable	*Code*	*Note*
x_1, nuclear defence	Score 1–21	Left–right-wing measure
x_2, unemployment *versus* inflation	Score 1–21	Left–right-wing measure
x_3, spending *versus* tax cuts	Score 1–21	Left–right-wing measure
x_4, nationalization *versus* privatization	Score 1–21	Left–right-wing measure
x_5, attitude towards Mrs Thatcher	Scalar 1–4	Very–not very effective
x_6, attitude towards Mr Foot or Mr Kinnock	Scalar 1–4	Very–not very effective
x_7, Conservative image 1	0, extreme; 1, others	
x_8, Labour image 1	0, extreme; 1, others	
x_9, Conservative image 2	0, united; 1, others	
x_{10}, Labour image 2	0, united; 1, others	

Table 1 lists the numbers of voters by their voting pattern for the Conservative Party in the three years.

As can be seen in Table 1, there is considerable dependence between responses in the three rounds of the panel study. For example, individuals voting Conservative on all three occasions made up 27.7% out of 729. Those voting consistently for or against the Conservative Party in both 1983 and 1986 make up 74.3% (202 + 14 + 37 + 325 + 57 + 119) among 1015 respondents appearing in the first two columns of Table 1. Similarly, the percentages of consistent votes are 63.3% in 1983 and 1987, and 66.2% in 1986 and 1987.

The explanatory variables of interest are given in Table 2. As measures of voters' fundamental values we use four scales, variables x_1–x_4 in Table 2, which measure voters' attitudes towards nuclear defence, unemployment (*versus* inflation), tax cuts (*versus* Government spending) and privatization (*versus* nationalization) (questions 24a, 28a, 31a and 36a in the 1983 BES). These were some of the central issues in the two general elections (see for example Butler and Kavanagh (1988), pages 216–221) and were ones on which most voters had relatively stable preferences. Attitudes towards these four issues were measured on 21-point scales: respondents were presented with two contrasting statements and asked to locate themselves at some point on a scale running from one statement to the other. The higher scores represent more 'right-wing' attitudes.

Two groups of variables reflect the more topical headline themes to which the voters will have been exposed over the course of the electoral cycle, namely evaluations of the political leaders (variables x_5 and x_6) and party images (variables x_7–x_{10}) in Table 2. Evaluations of the party leaders were asked on four-point scales. In 1983 respondents were asked

> 'on the whole how effective or ineffective do you think Mrs Thatcher is as a Prime Minister? And on the whole how effective or ineffective do you think Mr Foot would have been as a Prime Minister?'

Mrs Thatcher remained the leader of the Conservative Party (and of course Prime Minister) in 1987, but Mr Foot was replaced by Mr Kinnock as the Labour leader shortly after the 1983 election. In answering these questions on Prime Ministerial effectiveness, respondents were given four options — 'very effective', 'fairly effective', 'fairly ineffective' and 'very ineffective' together with the possibility of 'don't know'. For details about the categories from these questions see Heath *et al.* (1991). The percentages 'don't know' were 0.3% and 1.8% respectively and they were excluded from the analysis.

To ascertain their images of the parties, respondents were asked

> 'Moving now from the Party leaders to the Parties themselves. On the whole, would you describe the Conservative Party as extreme or moderate?'

There were four response codes, 'extreme', 'moderate', 'neither or both' and 'don't know'. For simplicity we have dichotomized these codes, contrasting extreme with the other three categories. A similar question was asked on perceptions of the parties as united or divided, and we followed the same procedure, dichotomizing the codes. We include the corresponding measures of the Labour Party's image.

The nature of the causal links between our different explanatory variables cannot of course be demonstrated by using our data alone. Political scientists have generally assumed that values are causally prior to images of parties and of leaders, and measures of values certainly can be shown to be considerably more stable over time than the image measures. (For a thorough discussion of the likely causal links between these variables see Bartle (1998).) It is also likely that images of parties and of leaders will be causally linked with each other, possibly in a reciprocal manner. Recognizing that the variables will be interrelated, our concern nevertheless is with the overall patterns for the different blocks of variables (i.e. with the patterns for the fundamental variables on the one hand and the image measures on the other hand) rather than the patterns for specific measures.

All our variables will contain amounts of measurement error. There is little information, however, for the BES about the size of this, especially in relation to the magnitude of changes over time. Woodhouse *et al.* (1996) discussed procedures for adjusting for measurement errors in multilevel models but we do not pursue this here, although we would expect that measurement errors will tend to weaken any associations in the data.

Our central hypothesis, then, is that the parameter estimates for variables x_1–x_4 would be relatively larger in the general election years of 1983 and 1987 whereas the estimates for variables x_5–x_{10} would have relatively greater influence on voting intention in the mid-term year of 1986.

3. Separate two-level models for each year

To obtain a feel for the data and a first view of the effect of the explanatory variables at each round of interviewing, we begin by fitting separate two-level logistic models to each year's data.

Denote by $\pi_{ij,t}$ the probability that individual i from the constituency j votes Conservative in year t. We model this as

Table 3. Separate models for 1983, 1986 and 1987†

	$\beta_{0,t}$	$\sigma^2_{u_{0,t}}$	*Extrabinomial parameter*
1983, $t=1$	−0.40 (0.079)	0.38 (0.094)	0.96 (0.036)
1986, $t=2$	−0.75 (0.076)	0.14 (0.086)	0.97 (0.046)
1987, $t=3$	−0.22 (0.078)	0.13 (0.091)	0.98 (0.051)

†Standard errors are given in parentheses.

$$\begin{gathered}\operatorname{logit}(\pi_{ij,t}) = \beta_{0,t} + \sum_{h=1}^{10} \beta_{h,t} x_{hij,t} + u_{0j,t}, \\ u_{0j,t} \sim N(0, \sigma^2_{u_{0,t}}).\end{gathered} \tag{1}$$

The observed (0, 1) response, at level 1, is $y_{ij,t} \sim \operatorname{bin}(1, \pi_{ij,t})$ with binomial variance $\pi_{ij,t}(1-\pi_{ij,t})$. The term $u_{0j,t}$ is the log-odds for Conservative voting in the jth constituency compared with the mean. This is a 'variance components' model without random coefficients. The assumption of binomial variation at level 1 can be tested by fitting 'extrabinomial parameter' σ^2_{e} so that the level 1 variance is $\sigma^2_{\mathrm{e}}\pi_{ij,t}(1-\pi_{ij,t})$. An estimate of σ^2_{e} close to 1.0 indicates close conformity to the Bernoulli distribution assumption (Goldstein (1995), chapter 7).

Table 3 gives the results for the simplest model, with only the intercept fitted. The extrabinomial variation is fitted with parameters here all close to 1, suggesting that the Bernoulli distribution is an adequate assumption for these data. The level 2 variances are slightly larger than their estimated standard errors, with the largest being for 1983.

We note that significance tests or confidence interval estimates for the variances based on these standard errors are very approximate. Those for the fixed parameters in these are more reliable and Wald tests based on the estimated covariance matrices of the fixed parameters will be used with subsequent models.

Table 4 gives the results when our 10 explanatory variables are added. Comparing the parameter estimates between years briefly, we see that parameters β_2 and β_4, associated with attitudes towards the unemployment and privatization issues, seem to show the cycle expected, being slightly smaller in 1986. But β_1, associated with attitudes towards nuclear defence, shows the reverse pattern.

Table 4. Parameter estimates for separate models for each year

Parameter	*Estimates for the following years:*		
	1983	*1986*	*1987*
β_0	−1.58 (0.34)	−0.73 (0.23)	−1.22 (0.36)
β_1	0.08 (0.01)	0.12 (0.02)	0.07 (0.03)
β_2	0.05 (0.01)	0.03 (0.01)	0.04 (0.02)
β_3	0.05 (0.02)	0.08 (0.02)	0.09 (0.03)
β_4	0.10 (0.01)	0.06 (0.02)	0.07 (0.03)
β_5	−1.06 (0.14)	−1.04 (0.15)	−1.33 (0.32)
β_6	0.55 (0.09)	0.57 (0.10)	0.79 (0.16)
β_7	1.13 (0.14)	1.13 (0.16)	1.25 (0.24)
β_8	−0.35 (0.14)	−0.39 (0.17)	−0.37 (0.25)
β_9	−0.67 (0.18)	−0.33 (0.17)	−0.76 (0.34)
β_{10}	0.27 (0.21)	0.06 (0.21)	0.64 (0.31)
$\sigma^2_{u_0}$, level 2	0.25 (0.10)	0.13 (0.08)	0.00
Extrabinomial, level 1	0.97 (0.04)	0.85 (0.04)	1.47 (0.03)

The two parameters β_5 and β_6, associated with attitudes towards Mrs Thatcher and Mr Foot or Mr Kinnock, show that the effect of leadership on voting intention was less in 1986 than it was on the vote in the following general election. This is the reverse of our hypothesis. Moreover, the parameters β_9 and β_{10}, associated with perceptions of the parties as united or divided, also follow an unexpected pattern, being smaller in 1986 than in the two election years. This first look at the data, then, does not suggest that headline information about party and leader images is more influential during the middle of the electoral cycle than it is at general election time.

The parameter estimates displayed in Table 4 do not allow us to draw any firm conclusions about the relative weight of the various variables at different stages of the electoral cycle. Fitting separate models to each wave of the panel data, as we do in Table 4, ignores the repeated information from the respondents who recorded votes on more than one occasion.

4. A three-level repeated measures model

To avoid the drawbacks of the separate models for each year, we can pool the data from each year into a single three-level repeated measures model. We treat year as the repetition at level 1 (indicated by t) nested within individuals (indicated by i), whereas individuals are nested within constituency j. Let z_t be the vector of indicator variables for $t = 1, 2, 3$ or 1983, 1986 and 1987 respectively,

$$\left.\begin{array}{ll} z_{1ij} = 1 & \text{if } t = 1983 \\ z_{2ij} = 1 & \text{if } t = 1986 \\ z_{3ij} = 1 & \text{if } t = 1987 \end{array}\right\} \text{and 0 otherwise.}$$

Since year is now level 1 our notation reflects this with t being the index for the first subscript. We shall use s_{tij} to denote the measurement of time (0, 3, 4) as a continuous score. We can write a model as follows for the probability of a positive response π_{tij}:

$$\operatorname{logit}(\pi_{tij}) = \sum_{t=1}^{3} \beta_{0,t} z_{tij} + \sum_{t=1}^{3} \sum_{h=1}^{10} \beta_{h,t} z_{tij} x_{h,tij} + \sum_{t=1}^{3} \nu_{tj} z_{tij} + u_{ij}, \tag{2}$$

$$u_{ij} = u_{0ij} + u_{1ij} s_{tij},$$

$$\nu_{tj} \sim N(0, \Omega_\nu), \qquad u_{ij} \sim N(0, \Omega_u),$$

$$\Omega_\nu = \begin{pmatrix} \sigma^2_{\nu 1} & & \\ \sigma_{\nu 12} & \sigma^2_{\nu 2} & \\ \sigma_{\nu 13} & \sigma_{\nu 23} & \sigma^2_{\nu 3} \end{pmatrix}, \qquad \Omega_u = \begin{pmatrix} \sigma^2_{u0} & \\ \sigma_{u01} & \sigma^2_{u1} \end{pmatrix},$$

where ν_{tj} are the residual terms at constituency level associated with the intercept for each year. Thus, for the jth constituency the marginal population mean (ignoring covariates) for Conservative voting for 1983, 1986 and 1987 on the logit scale is given by $\beta_{0,1} + \nu_{1j}$, $\beta_{0,2} + \nu_{2j}$ and $\beta_{0,3} + \nu_{3j}$ respectively. The level 2 variance is a quadratic function of time. As before we can model extrabinomial variation at level 1.

We first fit the simplest form of model (2). Only the MQL plus first-order approximation procedure provided converged estimates for this model. The results are given in Table 5.

Table 5. A variance component repeated measures model allowing extrabinomial variation

	Parameter	*Estimate (standard error)*
Fixed	$\beta_{0,1}$ (1983)	−0.39 (0.08)
	$\beta_{0,2}$ (1986)	−0.77 (0.07)
	$\beta_{0,3}$ (1987)	−0.30 (0.07)
Random, level 3	$\sigma^2_{\nu_1}$	0.34 (0.08)
	$\sigma_{\nu_{12}}$	0.28 (0.07)
	$\sigma^2_{\nu_2}$	0.20 (0.08)
	$\sigma_{\nu_{13}}$	0.28 (0.07)
	$\sigma_{\nu_{23}}$	0.21 (0.07)
	$\sigma^2_{\nu_3}$	0.20 (0.08)
Random, level 2	$\sigma^2_{u_0}$	2.37 (0.12)
	$\sigma_{u_{01}}$	0.0 (0.0)
	$\sigma^2_{u_1}$	0.0 (0.0)
Extrabinomial parameter		0.38 (0.01)

The predicted median proportions of Conservative votes in 1983, 1986 and 1987 are 0.40, 0.32 and 0.42 respectively from the fixed part of the model, which are close to the raw proportions of 0.41, 0.32 and 0.42. Note that these predictions are formed by taking the antilogits of the means on the logit scale. These transform to medians on the probability scale but in the present case with proportions close to 0.5 they will be very similar. In the general case the exact marginal distribution can be obtained via an approximation (Zeger *et al.*, 1988) or exactly via simulation (Goldstein (1995), chapter 5). The covariance terms at constituency level take up the dependence of the outcomes in the panel study.

However, there are major problems with this model. First, some correlations between years at constituency level are estimated to be greater than 1 and the covariance matrix is non-positive definite. One parameter at the individual level is estimated as 0 and the extrabinomial parameter is well below 1, suggesting that the assumption of binomial error for the model is not appropriate. As we noted earlier, a considerable proportion of the respondents voted in exactly the same way on all three occasions and it is reasonable to suppose that for a large minority their probabilities are 0 or 1. We shall not, therefore, consider this model further.

5. A multilevel multivariate logistic model

Using the same notation as in the case of the repeated measures model (2), a general multivariate logistic model for our data may be written

$$y_{tij} \sim \mathrm{bin}(1, \pi_{tij}),$$

$$\mathrm{logit}(\pi_{tij}) = \sum_{t=1}^{3} \beta_{0,t} z_{tij} + \sum_{t=1}^{3} \sum_{h=1}^{10} \beta_{h,t} z_{tij} x_{h,tij} + \sum_{t=1}^{3} \nu_{tj} z_{tij}. \tag{3}$$

We make the same assumptions as for the repeated measures model, except that there is no level 1 variation, but at level 2 we allow the binomial variates to covary. This is a convenient and efficient model for formulating a multivariate multilevel model (Goldstein, 1995).

At this level we estimate a covariance structure in which the diagonal terms are constrained to have binomial variance and the off-diagonal terms are estimated. From these estimates we

Table 6. A variance components multivariate model

	Parameter	*Estimate*†
Fixed	$\beta_{0,1}$ (1983)	−0.42 (0.08)
	$\beta_{0,2}$ (1986)	−0.81 (0.08)
	$\beta_{0,3}$ (1987)	−0.32 (0.08)
Random, level 3	$\sigma^2_{\nu_1}$	0.41 (0.10)
	$\sigma_{\nu_{12}}$	0.31 (0.08)
	$\sigma^2_{\nu_2}$	0.18 (0.09)
	$\sigma_{\nu_{13}}$	0.31 (0.08)
	$\sigma_{\nu_{23}}$	0.21 (0.08)
	$\sigma^2_{\nu_3}$	0.22 (0.09)
Random, level 2	$\sigma^2_{e_1}$	0.96 (0.04)
	$\sigma_{e_{12}}$	0.54 (0.03) (0.55)
	$\sigma^2_{e_2}$	1.00 (0.05)
	$\sigma_{e_{13}}$	0.62 (0.04) (0.64)
	$\sigma_{e_{23}}$	0.61 (0.04) (0.62)
	$\sigma^2_{e_3}$	0.98 (0.05)

†Standard errors are given in the first parentheses and biserial correlations in the second.

can obtain biserial covariances. We may also, as before, allow three extrabinomial variation parameters, one for each of the diagonal terms.

Estimates from the simplest multivariate model (without covariates) are given in Table 6. The predicted proportions of Conservative votes in years 1983, 1986 and 1987 are respectively 0.40, 0.31 and 0.42, which are comparable with the raw proportions of 0.41, 0.32 and 0.42. Comparing the results from the three separate models in Table 3 with those from the multivariate model in Table 6, both the fixed year effects and the estimated variances at the constituency and individual levels are reasonably close. We should not expect identical results from the two models because the model here estimates more random parameters at both constituency and voter levels to fit the dependence.

At the individual level the estimated extrabinomial parameters for the three binary responses are all close to 1, indicating that the assumption of a binomial error distribution for each time occasion is adequate. All three biserial covariances are large.

At the constituency level, the estimated variances for 1983, 1986 and 1987 are slightly larger than those from fitting the separate model (1) for each year, and the marginal distributions of the three sets of standardized constituency residuals are all close to normal. We note, however, that the level 3 estimated covariance matrix is non-positive definite. The estimation algorithm did not constrain this matrix to be positive definite, and given the relatively large standard errors we attribute the correlations which are slightly greater than 1.0 to sampling variability; the presence of a high intercorrelation at the constituency level is to be expected. It is also possible that the model is misspecified in some way but, as we shall see, this problem is resolved when further explanatory variables are fitted. Since there are on average 15 respondents per constituency the normality assumption for the estimated residuals seems reasonable.

We now elaborate the model by forming the interaction terms between the explanatory variables and the year indicators to fit the main effects for each year in the fixed part according to equation (3). To compare the effects of the same explanatory variables over the three years, we carry out joint tests (using approximate Wald statistics) for equality across

Table 7. Fixed part estimates from the multivariate model and tests for equality over occasions

Parameter	*Estimate (standard error) 1983*	*Estimate (standard error) 1986*	*Estimate (standard error) 1987*	χ_1^2, $\beta_{86}=\beta_{83}$	χ_1^2, $\beta_{86}=\beta_{87}$	χ_2^2, *joint test*
β_0	−1.27 (0.29)	−0.57 (0.20)	−0.98 (0.25)			
β_1	0.08 (0.01)	0.12 (0.02)	0.07 (0.02)	3.22	4.03†	4.77
β_2	0.05 (0.01)	0.03 (0.01)	0.03 (0.02)	1.12	0.09	2.01
β_3	0.05 (0.01)	0.07 (0.02)	0.08 (0.02)	1.21	0.13	2.30
β_4	0.09 (0.01)	0.04 (0.02)	0.05 (0.02)	5.41†	0.03	6.55†
β_5	−0.93 (0.12)	−0.92 (0.13)	−1.03 (0.21)	0.01	0.24	0.25
β_6	0.49 (0.08)	0.46 (0.09)	0.60 (0.10)	0.04	1.13	1.24
β_7	0.91 (0.12)	0.94 (0.14)	0.95 (0.17)	0.03	0.01	0.07
β_8	−0.39 (0.12)	−0.40 (0.15)	−0.34 (0.18)	0.01	0.08	0.08
β_9	−0.55 (0.15)	−0.24 (0.15)	−0.60 (0.24)	2.28	1.78	2.98
β_{10}	0.16 (0.27)	−0.10 (0.18)	0.57 (0.22)	0.66	6.03†	6.03†

†$p < 0.05$.

years, namely $\beta_{h,86} = \beta_{h,83}$ and $\beta_{h,86} = \beta_{h,87}$. The results are shown in the last three columns of Table 7.

Comparing the estimates for the variable effects in Table 7 with those in Table 4, we find as expected that many of the estimates are different. In particular some of the larger estimates, e.g. β_5, β_7 and β_9, have been reduced in size. However, the broad pattern over the three occasions does not change.

The joint test for equality across years produces significant results in the case of only two variables: attitudes towards privatization (x_4) and the image of the Labour Party as united or divided (x_{10}). In both cases the pattern is for the parameter estimate to be weaker in 1986 than in the two election years of 1983 and 1987, although the separate tests show no significant difference between 1983 and 1986 for the effect of x_4, nor between 1987 and 1986 for that of x_{10}.

Overall, this does not support our theory that fundamental variables, such as attitudes towards dominant issues, are more important at election time whereas headline topics such as party and leader images are more important during the middle of the cycle. Possibly this is because the non-election round of interviews was conducted rather too late in the electoral cycle, being held in the autumn of 1986, less than 12 months before the June 1987 election. By the autumn, the Conservatives had already recovered their popularity in the opinion polls and the panel study did not therefore really capture the phase of mid-term disillusion with the Government.

In Table 8 we list the random parameter estimates, although we shall not explore them further in this paper.

The results in Table 8 show that, once the attitude and image variables have been included in the fixed part of the model, the estimated variations at level 3 are much reduced from those in Table 6 with zero variation estimated for 1987 and a very small variance for 1986. It appears that the Conservative vote was constant among constituencies over the electoral cycle with much variation being accounted for by these explanatory variables, especially in the 1986 and 1987 elections. We also note that there is a considerable reduction in the between-year correlation for voters, and some underdispersion at voter level also. Constraining the model to fit binomial variation at the lowest level does not, however, appreciably alter any of the other parameter estimates.

Finally, we should point out some limitations of this panel study. First, the non-election round of interviews was conducted rather late so the panel did not really capture the phase of

Table 8. Random parameter estimates from the multivariate model

Level		*Random effects*	*Correlation*
3	$\sigma^2_{\nu_1}$	0.25 (0.09)	
3	$\sigma_{\nu_{12}}$	0.15 (0.07)	
3	$\sigma^2_{\nu_2}$	0.11 (0.09)	
3	$\sigma_{\nu_{13}}$	0.0	
3	$\sigma_{\nu_{23}}$	0.0	
3	$\sigma^2_{\nu_3}$	0.0	
2	$\sigma^2_{e_1}$	0.86 (0.03)	
2	$\sigma_{e_{12}}$	0.25 (0.03)	0.30
2	$\sigma^2_{e_2}$	0.80 (0.04)	
2	$\sigma_{e_{13}}$	0.26 (0.03)	0.29
2	$\sigma_{e_{23}}$	0.25 (0.03)	0.29
2	$\sigma^2_{e_3}$	0.92 (0.05)	

Table 9. Fixed effect estimates and test for trend of cycle

Year	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}
1983	0.10	0.05	0.09	0.13	−0.95	0.71	1.09	−0.86	−1.07	0.71
1986	0.13	0.04	0.07	0.04	−1.13	0.89	0.87	−1.00	−0.35	0.90
1987	0.07	0.05	0.09	0.07	−1.20	0.49	0.49	−0.94	−0.59	0.28
χ^2_1	12.0†	0.51	0.91	21.4†	0.15	0.59	0.39	0.68	10.7†	6.3‡

†$p < 0.01$.
‡$p < 0.05$.

the mid-term disillusion with the Government. Secondly, the data set is relatively small with 3357 responses from 1613 respondents in 112 constituencies. Thirdly, most of the variables were constructed to reflect the difference between two extremes (left–right wing) and two parties (Conservative and Labour), whereas our model contrasts Conservative and all others. For this reason we have also fitted the model to responses of Conservative *versus* Labour only (2031 responses from 1150 voters in 112 constituencies) and conducted a trend test for the election cycle on each variable. The contrast coefficients vector for the first four variables is (1, −2, 1) and that for the last six is (−1, 2, −1) for the assumed pattern. This gave us the findings in Table 9.

We see that variables 2–4 show the expected cycle, and also that the preference variables x_6, x_8 and x_{10} show the expected pattern. Variables x_1 and x_9 still show the reverse pattern.

6. Discussion

Using three-level multivariate logistic models, we can test the theory of election cycles based on one set of panel data. The overall findings lead to some general conclusions around our substantive theory. First, there is some evidence to support our assumption about the effects of voters' preferences and ideologies on their voting behaviour over the electoral cycle. Secondly, there is evidence in our study suggesting that there is more context effect in the general election of 1983 than in the other two years. The main effects of the covariates have explained most of the constituency effect in the three years. Substantively it is not important to explore the small random effect that is left for 1983 only.

Whereas a three-level repeated measures logistic model seems to provide a natural way to model these data, our study has demonstrated that the level 1 variation is seriously underdispersed because some individuals have a constant response and it is therefore not generally suitable for such data.

The multilevel multivariate logistic model assumes binomial error at each occasion with the covariance structure at voter level estimated to account for the dependence between the repeated outcomes. It has the same advantages as the repeated measures model in terms of the efficiency from pooling all the data in one model. The model's predictions for the overall probability of voting show a reasonable agreement with the raw probabilities. The estimated variance among constituencies for each year is similar to that from the marginal models fitted to each year separately, and the binomial assumption holds for the lowest level error distribution by year. It is also possible to generalize the multivariate model for the general repeated measures case with any number of occasions, but this will involve setting up an explicit model for the autocorrelation structure, and work on this is currently under way (Barbosa and Goldstein, 1999). Further work is also under way using models including more general variables for fundamental issues and extending our model to fit multiple-category responses over repeated occasions.

Another advantage of the multilevel framework is that the constituency level residuals can be further modelled. For example constituencies with different political characteristics may vary in their level of tactical voting from year to year. Let variable d_j indicate the distance from contention of the Conservative Party in the jth constituency. This is defined as the percentage difference between the vote for the Conservatives and that for the winner in the given constituency at the previous general election. We could then, for example, model its effect on the support for the party in each of the three years across constituencies as

$$\nu_j = (\nu_{1j} + \gamma_{1j} d_j) z_{1ij} + (\nu_{2j} + \gamma_{2j} d_j) z_{2ij} + (\nu_{3j} + \gamma_{3j} d_j) z_{3ij}. \tag{4}$$

This model would be straightforward to fit using the techniques of this paper. Nevertheless, because of the way that our variables were constructed to discriminate between Conservative and Labour, inferences about tactical voting need to be handled with caution.

Acknowledgements

This work was carried out as part of the Multilevel Models Project funded by the Economic and Social Research Council (UK) under the programme for the Analysis of Large and Complex Datasets. We are very grateful to the referees for comments on the earlier drafts.

References

Barbosa, M. F. and Goldstein, H. (1999) Discrete response multilevel models for repeated measures; an application to voting intentions data. To be published.

Bartle, J. (1998) Left-right position matters, but does social class?: causal models of the 1992 British General Election. *Br. J. Polit. Sci.*, **28**, 501–529.

Breslow, N. E. and Clayton, D. G. (1993) Approximate inference in generalised linear models. *J. Am. Statist. Ass.*, **88**, 9–25.

Butler, D. and Kavanagh, D. (1988) *The British General Election of 1987*. London: Macmillan.

Cox, D. R. (1972) The analysis of multivariate binary data. *Appl. Statist.*, **21**, 113–120.

Diggle, P. J., Liang, K.-Y. and Zeger, S. L. (1994) *Analysis of Longitudinal Data*. Oxford: Clarendon.

Gelman, A. and King, G. (1993) Why are American Presidential election campaign polls so variable when votes are so predictable? *Br. J. Polit. Sci.*, **23**, 409–451.
Goldstein, H. (1986) Multilevel mixed linear model analysis using iterative generalised least squares. *Biometrika*, **73**, 43–56.
———(1991) Nonlinear multilevel models with an application to discrete response data. *Biometrika*, **78**, 45–51.
———(1995) *Multilevel Statistical Models*. London: Arnold.
Goldstein, H. and Rasbash, J. (1996) Improved approximations for multilevel models with binary responses. *J. R. Statist. Soc.* A, **159**, 505–513.
Heath, A. F., Jowell, R. M. and Curtice, J. K. (1985) *How Britain Votes*. Oxford: Pergamon.
Heath, A. F., Jowell, R. M., Curtice, J. K., Evans, G., Field, J. and Witherspoon, S. (1991) *Understanding Political Change: the British Voter 1964-1987*. Oxford: Pergamon.
Heath, A., Yang, M. and Goldstein, H. (1996) Multilevel analysis of the changing relationship between class and party in Britain 1964-1992. *Qual. Quant.*, **30**, 389–404.
Korn, E. L. and Whittemore, A. S. (1979) Methods for analysing panel studies of acute health effects of air pollution. *Biometrics*, **35**, 795–802.
McLean, I., Heath, A. F. and Taylor, B. (1996) Were the 1994 Euro- and local elections in Britain really second-order? In *British Elections and Parties Yearbook 1996* (eds D. M. Farrell, D. Broughton, D. Denver and J. Fisher), pp. 1–20. London: Cass.
Miller, W. L. and Mackie, T. (1973) The electoral cycle and the asymmetry of government and opposition popularity. *Polit. Stud.*, **21**, 263–279.
Miller, W. L., Tagg, S. and Britton, K. (1986) Partisanship and party preferences in government and opposition: the mid-term perspective. *Elect. Stud.*, **5**, 31–46.
Rasbash, J., Browne, W., Goldstein, H., Yang, M., Plewis, I., Healy, M., Woodhouse, G. and Draper, D. (1999) *A User's Guide to MLwiN*, 2nd edn. London: Institute of Education.
Reif, K. (1984) National electoral cycles and European elections, 1979 and 1984. *Elect. Stud.*, **3**, 244–255.
Stray, S. and Silver, M. (1983) Government popularity, by-elections and cycles. *Parl. Aff.*, **36**, 49–55.
Woodhouse, G., Yang, M., Goldstein, H. and Rasbash, J. (1996) Adjusting for measurement error in multilevel analysis. *J. R. Statist. Soc.* A, **159**, 201–212.
Zeger, S. L., Liang, K. and Albert, P. (1988) Models for longitudinal data: a generalised estimating equation approach. *Biometrics*, **44**, 1049–1060.