MANCHESTER
1824

The University
of Manchester

Cathie Marsh Centre for Census and Survey Research

Modeling multiparty elections, preference classes and strategic voting.

CCSR Working Paper 2006-01
Ed Fieldhouse, Andrew Pickles, Nick Shryane, Jerry Johnson, Kingsley Purdam

We examine the role of strategic motivations in mediating the relationship between underlying political preferences and vote choice, in a multi-party, single member, simple plurality system. Political preference data from the British Election Panel Survey, 1997- 2001, were modeled with mixed multinomial logit models.
.

www.ccsr.ac.uk

Modeling multiparty elections, preference classes and strategic voting.

Dr. Edward Fieldhouse, Director, e.fieldhouse@manchester.ac.uk, +44 (0)161 275 4721

Professor Andrew Pickles, andrew.pickles@manchester.ac.uk, +44 (0)161 275 5204

Dr. Nick Shryane, n.shryane@manchester.ac.uk, +44 (0)161 275 4738

Jerry Johnson, jerry.johnson@stud.manchester.ac.uk, +44 (0)161 275 7025

Dr. Kingsley Purdam, kingsley.purdam@manchester.ac.uk, +44 (0)161 275 4719

Centre for Census and Survey Research, Crawford House, University Precinct,
University of Manchester, Oxford Road, Manchester, M13 9QS, UK.
Fax. +44 (0)161 275 4722

Abstract

We examine the role of strategic motivations in mediating the relationship between underlying political preferences and vote choice, in a multi-party, single member, simple plurality system. Political preference data from the British Election Panel Survey, 1997-2001, were modeled with mixed multinomial logit models. Latent variables were used to model the stable party political traits underlying observed preferences, allowing correlation between choices and so avoiding the restrictive assumption of independence from irrelevant alternatives. The assumption of normally-distributed latent political traits was found to be violated, so a non-parametric approach was used to specify a discrete distribution of latent classes. In addition to vote data, ranked approval ratings were used to help identify the models and to better characterize the underlying political preferences in the presence of insincere voting, which disproportionately affects 'third' parties. From these models we estimate that approximately 9% of votes cast may have been affected by strategic factors.

Strategic, non-sincere or tactical voting is widely recognised as being increasingly influential in modern elections, especially in single member plurality systems where there are more than two effective competitors (e.g. Great Britain). However, there is no clear consensus on how strategic voting should be measured, nor on the extent to which subjective notions of strategic voting are consistent with the objective conditions under which such behaviour should (rationally) take place. Blais and Nadeau (1996), and Alvarez and Nagler (2000) identify three approaches to identifying strategic voting based on (i) using aggregate data (e.g. Spafford, 1972; Johnston & Pattie, 1991); (ii) self-reported strategic voting (e.g. Evans and Heath, 1993; Niemi et al, 1993), and; (iii) divergence between stated preferences and vote (e.g. Black, 1978; Blais and Nadeau, 1996). Recognising the potential weaknesses in each of these approaches, like Alavarez and Nagler we adopt an approach based on the modeling of discrepancies between stated preferences and reported behavior, whilst taking into account the objective strategic context (c.f. Cain, 1978). The modelling approach, which utilizes repeated response data to identify latent preference classes in mixed multinomial logit models, is also substantially different and (as we show) more likely to provide reliable estimates. We draw upon data from the British Election Panel Study for the 1997 and 2001 General Elections in order to examine the extent of 'non-preference voting' broadly defined, and strategic voting, both subjectively and objectively defined. The approach both allows us to better represent preferences and voting for third or minor parties in a multiparty system, in this instance the British Liberal Democrats, and provides a more general framework for modelling party preference and voting behaviour.

It has been argued previously that preference and vote are not equivalent and that voters are more likely to betray their preferences where the objective conditions dictate (Black, 1978; Cain, 1978). In other words voting contrary to one's stated party preferences is related to the electoral context, or more precisely in the U.K., to the strategic situation in the constituency (electoral district), as this will shape the likelihood that a vote for a preferred candidate will be influential in the final outcome of the particular election or district. The rationale for this is straightforward: should a voter perceive a vote cast for their preferred candidate to be most likely wasted, that voter may opt to vote for another party. This is equally true in presidential elections where strategic voting may depend on the viability of the preferred candidate (e.g. Abramson et al, 1988). This is consistent with a rational choice approach and decision theory: McKelvey and Ordeshook and others have shown that a rational voter may vote against his or her most preferred party where it has little or no chance of winning in a multiparty election (McKelvey and Ordeshook, 1972; Cain, 1978). This can also be considered 'Duvergian' tactical voting (Fisher, 2004) as it is consistent with Duverger's law which states that third or lower-placed parties are deserted by rational voters in a single member simple plurality system (Riker, 1982, after Duverger, 1954). We would expect that strategic voting is more likely to affect smaller parties, as they are less likely to be considered viable, and parties in the political centre since (if preferences are single-peaked) they are more likely to be the second preference of voters on both the left and the right. Thus, in the case of Britain it is therefore unsurprising that strategic voting is particularly evident in the case of the Liberal Democrats (Spafford, 1972; Russell & Fieldhouse, 2004).

However, research into recent British elections has tended to focus on the vote share between the two major parties, Labour (on the center-left) and Conservative (on the center-right). The third-largest party, Liberal Democrat (centrist), has often been ignored or cast as an influence on the other parties rather than as a straightforward alternative in its own right (e.g. Sanders, 1996). This is despite support for the Liberal Democrats being quite substantial, with around 20% of the vote share in general elections since the 1980s.

Many of the simplifications taken when modeling multiparty elections have been driven by the availability of methodological tools for binary choices, where much analysis has been between just the top two parties or between a target party *versus* the remainder. This approach does not deal with the actual choices that voters face, however, and has been found to misestimate vote share between the larger parties in part because of the influence of the overlooked smaller party (Heath et al, 1991). More realistic modeling approaches must deal with the polytomous choice presented in multiparty elections.

Multinomial logit (MNL) allows for the modeling of discrete choices with more than two alternative parties (e.g. Whitten & Palmer, 1996). However, it has been criticized because of its restrictive assumption of "independence from irrelevant alternatives" (IIA). With IIA the relative probability of choosing between any pair of alternatives must not depend on the presence (or absence) of other alternatives in the choice set. In effect, this assumption implies that voters must not view any parties as being clustered together, in the sense that the parties in the cluster could be seen as substitutes. This is patently not the case in British party politics, where the Liberal Democrats are currently seen as a closer substitute for the Labour Party than for the Conservative Party (Russell and Fieldhouse, 2004).

Other authors have used multinomial probit (MNP) models, which allow for (limited) specification of a covariance structure for the disturbances and so do not impose the IIA assumption (e.g. Alvarez & Nagler, 1995; Alvarez, Nagler & Bowler, 2000). A recent review (Dow & Endersby, 2004) found little difference between MNL and MNP estimates when modelling the same UK multiparty election data, however.

We use mixed multinomial logit (MMNL) models for our analyses. The advantage of MMNL is its flexibility; any discrete choice model consistent with random utility maximization can be approximated by a MMNL model (McFadden & Train, 2000). This is not the case for the standard MNP, where, for example, non-normally distributed random terms cannot be accommodated (Hensher & Greene, 2003). This is of significance later in this paper, where, because of non-normally distributed latent party preferences, an approximate non-parametric maximum likelihood estimator (NPMLE) solution is used to estimate latent classes of political preference.

Preference and vote in the 1997 & 2001 general elections in England

Data from two waves of the British Election Panel Study 1997-2001 (BEPS-2; Thomson et al 1999) were used. These waves, part of the 30-year, ongoing British Election Studies, were conducted shortly after the 1997 and 2001 general elections. Although the survey was also conducted in Wales and Scotland, only respondents living in England were utilized as they represented the largest group with the same choice set of three parties in each constituency – Conservative, Labour and Liberal Democrat.

There were 2,551 respondents in the panel in the 1997 wave and 1,709 remaining in the 2001 wave. Only data from waves where individuals had a clear party of first preference and reported voting for one of the three major parties were retained. (Support for parties other than the main three was very limited; only 69 respondents reported voting for a different party, and an additional 17 said that they supported other parties even though they voted for a main party.) Further, 75 respondents with missing covariates (discussed later) were discarded. This resulted in 1,914 respondents contributing matched preferences and votes for 2,892 voting occasions.

The BEPS-2 did not require respondents to explicitly rank-order the main parties in terms of preference. Instead, a ranking was constructed based upon party approval ratings and party identification. Respondents were asked to rate their current approval of each of the main parties on a 5-point scale from “strongly disapprove” to “strongly approve”. These ratings were then ranked to give a preference ordering. Because of the limited response options with a 5-point scale tied rankings were common. This was most problematic where first preference was tied as no clearly preferred party could be identified. Full rankings were only available for around 67% of all sets of ratings.

Although tied ranks may have indicated genuine indifference on the part of the voters, other questions in the BEPS-2 survey proved useful in differentiating first and second choice for most respondents with tied preferences. Fisher (2004) argued for using reported vote to make this distinction, unless the voter had reported voting strategically.

However, there are two problems with this: first there is ambiguity associated with self-report measures of strategic motivation (e.g. see Blais and Nadeau, 1996; Niemi et al, 1993); second we require a measure of preference *independent* of vote, as vote choice is included elsewhere in the model. The BEPS/BES measure of party identification was considered appropriate for this purpose, as it has been shown to tap the underlying concept of current party preference or 'valenced partisanship' in the U.K. (Brynin and Sanders, 1997; Clarke et al. 2004), as opposed to the long-standing affiliation it has been found to represent in the USA (e.g. Campbell et al. 1964; Green et al, 2004). In virtually all cases party identification was consistent with the respondent's approval ratings. Where approval ratings were tied for first place therefore, party identification was used to decide the preference ranking. Of the final sample, a full rank ordering of the parties was derived for 76%, with 24% having a clear first preference but tied preference for second place.

The structure of the ranked party preferences is illustrated graphically in Figure 1, in a variation on the form of a permutahedron (Schulman, 1979; Zhang, 2004). Each full rank-ordering of the three parties (3! = 6) was assigned an axis and the axes arranged like the spokes of a wheel with their origins at the hub. The relative placements of the axes reflect the distances between the rank orderings (distance as defined by Kendall's (1938) τ correlation coefficient), so that the closest are adjacent and the most distant ($\tau = -1$) are opposite one another across the origin. The three pairs of opposite axes have inverse rank orderings and so represent 'major' axes of relative preference between pairs of parties. In Figure 1 the Labour – Conservative axis is labeled (1), the Liberal Democrat – Conservative axis is (2) and the Labour – Liberal Democrat axis is labeled (3).

----Figure 1 about here----

The six wedge-shaped sectors between the axes each have the same party ranked either first or last on their two defining axes, and so each sector represents valence towards or against a party. Each sector is labeled with the relevant party name and a plus sign to indicate relative preference and a minus sign to indicate relative aversion. For example, the bottom-right sector is labeled "LD+", because the Liberal Democrats are ranked first on both of the axes that define the sector.

For each wave separately the percentage of respondents with each ranking was plotted on the appropriate axis and the points joined to form a polygon. Of all the rankings 24% had tied 2nd preferences and were therefore identified uniquely only to a sector, not to an axis. These preferences were allocated to axes in proportion to the ratio of fully ranked preferences between the axes that constituted the sector[1]. Figure 1 shows that the pattern of preferences changed little from 1997 to 2001, with slight gains for the Conservatives and the Liberal Democrats at the expense of Labour.

Considering first the major axes, the Labour-Conservative axis (1) was the most significant, containing 70% of preferences overall, split 2:1 between Labour and Conservative preferrers, respectively. The Liberal Democrat-Conservative axis (2) accounted for 24% of the preferences, again split approximately 2:1 against the Conservatives. The Labour-Liberal Democrat axis was relatively trivial, with only 6% of respondents with preferences on this axis.

[1] For example, of all the 1997 rankings, 12% were "Con, (Lab, LD)", i.e. Conservatives ranked first and the other two parties tied. These preferences were allocated to "Con, LD, Lab" and "Con, Lab, LD" in the ratio of full rankings between these axes, 7:3 in this case. That meant 8.4% were plotted on the "Con, LD, Lab" axis and 3.6% on the "Con, Lab, LD" axis.

In terms of sectors, the Lab+ sector contained 49% of all preferences, Con+ 32% and LD+ 19%. However, the Con- sector was the largest overall, accounting for 62% of preferences. Combining both Con+ and Con- sectors together accounts for fully 94% of preferences; this, coupled with the limited extent of the Labour-Liberal Democrat axis, reflects the notion that the Liberal Democrats and Labour are not seen as independent alternatives but are part of an anti-Conservative bloc.

In terms of a single ideological Left-Right political dimension, the Conservatives are posited to be on the right and Labour on the left, with the Liberal Democrats in between (although closer to Labour, as suggested above). The above rankings provide good evidence in favor of this predominately uni-dimensional interpretation of the preferences; the rankings that are not single-peaked on this continuum ("Lab, Con, LD" and "Con, Lab, LD", i.e. the LD- sector) account for only about 7% of the rankings overall.

In the sectors with positive party valence the "V" boxes show the percentage of respondents who reported voting for their preferred party. Over 90% of both Conservative and Labour preferrers said that they voted with their party preference, compared to only 78% of Liberal Democrat preferrers, demonstrating the increase in non-preference voting for preferrers of the "third" party. Overall, across both waves, Labour attracted 47% of the reported vote, the Conservatives 33% and the Liberal Democrats 20%. Given that Liberal Democrat supporters were much less likely to vote for their preferred party than were other party preferrers, the fact that their vote share was slightly greater than their overall level of first preferences suggests that the Liberal Democrats gained votes from other parties' supporters at least as much as they lost votes to other parties from their own supporters (see Russell and Fieldhouse, 2004).

Modeling framework

The structure of these party preferences and their relationship to voting was specified as a MMNL model. In the multinomial models the three parties represent a choice set with one of the choices as a reference category. The probability of choosing (or giving a higher ranking to) the other parties is then estimated. Derivation of the MMNL model is illustrated next.

In a random utility model, U^a (here, a = 1, 2, 3), the subjective value of alternative a, i.e. utility, is modelled as being comprised of two parts: V^a, the measured characteristics of the chooser or choice alternative, e.g. age of voter, age of candidate, and ε^a, a random component representing unmeasured idiosyncrasies,

$$U^a = V^a + \varepsilon^a .$$

Alternative 1 is then chosen over alternative 2 if $U^1 > U^2$, or equivalently,

$$U^1 > U^2 = V^1 - V^2 + (\varepsilon^1 - \varepsilon^2) > 0 .$$

If ε^a has an Extreme Value type I distribution then the differences $(\varepsilon^1 - \varepsilon^2)$ have a logistic distribution (McFadden, 1974), and the probability that U^1 exceeds U^2 is,

$$\Pr(U^1 > U^2) = \frac{\exp(V^1)}{\exp(V^1) + \exp(V^2)} .$$

If V^a is parameterized as a linear combination of subject-specific covariates $\mathbf{x}$, and the coefficients for the third, reference alternative are set to zero (for identification), the familiar multinomial logit model results, e.g. for the probability of choosing alternative 1 compared to the reference alternative (here, alternative 3):

$$\Pr(choice=1)=\frac{\exp(\boldsymbol{\beta}^1\mathbf{x})}{1+\exp(\boldsymbol{\beta}^1\mathbf{x})+\exp(\boldsymbol{\beta}^2\mathbf{x})}.$$

This specification is dependent on the assumption that $(\varepsilon^1-\varepsilon^3)$ and $(\varepsilon^2-\varepsilon^3)$ are independently distributed (the IIA assumption). However correlation among the random terms ε^a is in practice likely. The correlations can be modelled by introducing shared random effects u^a,

$$U^a=V^a+u^a+\varepsilon^a.$$

The differences between the random effects for each category compared to the reference are denoted η^a,

$$\eta^1=u^1-u^3$$

$$\eta^2=u^2-u^3.$$

They reflect the propensity to choose an alternative compared to the reference after the measured covariates have been accounted for. The probability of choosing alternative 1 then becomes

$$\Pr(choice=1)=\iint_{\eta^1\eta^2}\frac{\exp(\boldsymbol{\beta}^1\mathbf{x}+\eta^1)}{1+\exp(\boldsymbol{\beta}^1\mathbf{x}+\eta^1)+\exp(\boldsymbol{\beta}^2\mathbf{x}+\eta^2)}dG(\eta^1\eta^2)$$

for some distribution $G(\eta^1\eta^2)$ of the latent variables. Commonly the latent variables are assumed multivariate normal but other forms may be empirically more suitable. In general, the latent variables that give rise to the correlations among choices can be poorly identified if only first-choice information is used (Skrondal & Rabe-Hesketh, 2003). This limitation can be overcome by using ranked preferences instead of first-choices.

The Luce model for ranked preferences is a direct extension of the random utility derivation of the multinomial choice model (Luce, 1959). With three alternatives the first choice probabilities are as for the original model; second choice probabilities, conditional on the first choice, are given by the same multinomial form, but with the first-choice excluded from the choice set. For example, with three alternatives, the probability that alternative 1 will be ranked first, followed by alternative 2 second (with the final choice redundant) is

$$\Pr(1st\ choice = 1, 2nd\ choice = 2) = \frac{\exp(\boldsymbol{\beta}^1\mathbf{x})}{1+\exp(\boldsymbol{\beta}^1\mathbf{x})+\exp(\boldsymbol{\beta}^2\mathbf{x})} \cdot \frac{\exp(\boldsymbol{\beta}^2\mathbf{x})}{1+\exp(\boldsymbol{\beta}^2\mathbf{x})}$$

Allowing for correlations among utilities with latent variables gives

$$\Pr(1st\ choice = 1, 2nd\ choice = 2) =$$

$$\iint\limits_{\eta^1\eta^2} \frac{\exp(\boldsymbol{\beta}^1\mathbf{x}+\eta^1)}{1+\exp(\boldsymbol{\beta}^1\mathbf{x}+\eta^1)+\exp(\boldsymbol{\beta}^2\mathbf{x}+\eta^2)} \cdot \frac{\exp(\boldsymbol{\beta}^2\mathbf{x}+\eta^2)}{1+\exp(\boldsymbol{\beta}^2\mathbf{x}+\eta^2)} dG(\eta^1\eta^2)$$

The case where the first choice is known but there is a tied ranking between 2nd and 3rd choices simplifies to the specification for first choices given earlier.

Finally, the fixed part of the model V^a can accommodate covariates that vary over observations and alternatives:

$$V_j^{\,a} = \mathbf{x}'_j\mathbf{g}^a + \mathbf{x}'^{a}_j\mathbf{b}$$

where $\mathbf{x}'_j$ is a vector of covariates that varies over respondents *j* but not alternatives (e.g. the sex of the voter), whereas $\mathbf{x}'^{a}_{j}$ may also vary over alternatives (e.g. the sex of the candidate). The corresponding coefficient vectors are $\mathbf{g}^{a}$ and $\mathbf{b}$. The coefficient $\mathbf{b}$ that multiplies the alternative-specific covariates is assumed to be the same for all utilities and so represents the linear effect between utility and the covariates. In econometrics this would represent, say, the effect of cost, which may vary over alternatives but is assumed to have the same magnitude of effect for all choices. However, for some effects this may not hold true, in which case interactions between these variables and dummy variables for the alternatives in $\mathbf{x}'^{a}_{j}$ can be specified, allowing the alternative-specific predictors to vary in their effects across alternatives.

Models of preference and vote

First, simple models of political inclination based upon ranked preference and vote were estimated. Two equations were specified quantifying the probability of preferring either the Labour party ($a = 1$) or the Liberal Democrats ($a = 2$) in comparison to the Conservative party ($a = 3$). The data were from two occasions *i* ($i = 1, 2$; 1997 and 2001), nested within individuals, *j* ($j = 1 \ldots J$).

Considering first the fixed part of the model *V*, for each comparison this contained two election-specific constants (1997 and 2001) of the overall relative preference for each party at each election based on both ranked preferences and voting responses. A dummy variable, VOTE, equal to 1 for voting responses and 0 otherwise, was also defined to provide a contrast with overall preference. This produced four occasion-specific covariates $\mathbf{x}_{ij}$;

$$V_{ij}^{\ a} = x_{1997\,j}\, g_1^{\,a} + x_{1997*VOTE\,j}\, g_2^{\,a} + x_{2001\,j}\, g_3^{\,a} + x_{2001*VOTE\,j}\, g_4^{\,a}$$

The random parts of the model, η_j^a, were considered as party-political traits that varied across voters but not elections. These random effects were used to model the unobserved heterogeneity between voters, and represented unobserved party political traits for Labour (η_j^1) and the Liberal Democrats (η_j^2) in preference to the Conservative party. The correlation between these latent variables, ψ_{η^a}, was also freely estimated (as per Skrondal & Rabe-Hesketh, 2003), thus yielding correlated, alternative-specific random intercepts at the voter level. This model was named Model 2, in reference to the number of latent variables.

The above model allowed latent Labour vs. Conservative traits to vary independently from latent Liberal Democrat vs. Conservative traits. The structure of ranked preferences in Figure 1 suggested that the party preferences might have been uni-dimensional, however, with the distinction between Labour and Liberal Democrat preferences in relation to the Conservatives being merely one of degree. For this reason an additional, single-factor, model was fitted. For this model the fixed part was identical to that above, but the random part featured only one latent variable η_j, which can be conceived as representing unobserved heterogeneity on a latent Left-Right (or pro-, anti-Conservative) dimension. To allow separate effects of this variable on observed Labour and Liberal Democrat preferences and votes the latent variable was multiplied by an alternative-specific factor loading λ_j^a. To allow model identification the scale of the factor was anchored by fixing the first factor loading to one, in this case the loading for the Labour responses, $\lambda_j^1 = 1$, whilst allowing the Liberal Democrat vs. Conservative loading λ_j^2 to be estimated freely. This model was named Model 1.

Finally, the assertion that the clustering of Labour and Liberal Democrat preferences in relation to the Conservatives' was actually in violation of the IIA assumption was tested by also fitting a standard MNL model without latent variables to the data. This model comprised the fixed part of the above models but with no random component η_j, and was named Model 0. In models 1 and 2 the latent variables were assumed normally distributed.

Figure 2 shows a path diagram of Model 2; the uni-dimensional Model 1 is the same but with only a single latent variable; Model 0 is the same but without latent variables.

--- Figure 2 about here

These and subsequent models were estimated in `gllamm` (Rabe-Hesketh, Pickles & Skrondal, 2001) using maximum likelihood, with latent variable distributions approximated by Guass-Hermite quadrature (Stroud, 1971). Robust standard errors are reported, adjusted for the clustering of individuals within electoral districts. (`gllamm`, freely available from www.gllamm.org, runs in Stata (StataCorp, 2003) and can estimate a wide range of generalized linear latent and mixed models.)

Table 1 shows the log-likelihoods for the models, as well as the latent variable estimates (for models 2 and 1). Because the models are nested likelihood ratio tests were used to assess relative model fit, and revealed highly significant differences in fit among all models.

----Table 1 about here----

First, it can be seen that the conventional MNL Model 0 provided by far the worst fit to these data. This model took no account of the unobserved heterogeneity among preferences across voters. Second, both models 2 and 1 showed that there is less “distance” between latent Liberal Democrat and Conservative preferences compared to Labour and Conservative preferences, which is as would be expected if Labour and Conservative exist at opposite end of the left-right political spectrum with the Liberal Democrats in between. In Model 1 this was demonstrated by the 0.66 factor loading for Liberal Democrat preferences, approximately two-thirds that of the reference loading for Labour. Similarly, in Model 2 the variance of the Labour latent variable was more than double the variance of the Liberal Democrat latent variable, i.e. greater variability between Labour and Conservative preferences than between Liberal Democrat and Conservative preferences. Third, the correlation between latent variables in Model 2 was .81 – a strong but far from perfect relationship, suggesting that the single dimension of latent preference specified in Model 1 was an oversimplification. Indeed, this was borne out by the model log-likelihoods, which showed that Model 1 provided a significantly worse fit to the data than Model 2.

Quadrature issues

The models reported in Table 1 were estimated under the assumption that the latent party political traits were continuous variables with normal distributions. These distributions are approximated in `gllamm` by summing the likelihood contributions from a limited number of discrete points, i.e. Gauss-Hermite quadrature (see Stroud, 1971). In general the more points that are used the better is the approximation and the more accurate is the characterization of the latent variable. However, numerical integration of this kind is computationally demanding and as more points are used the estimation time increases rapidly, especially when fitting more than one latent variable. A practical way around this problem is to take the fewest number of points that will generate acceptable results. Acceptability is checked by estimating the model with greater numbers of quadrature points and if the coefficients are found not to change much the solution is accepted. Bhat (2001) suggests that 10 quadrature points per latent variable is a viable minimum number to use when estimating models with two latent dimensions (as here); Longford (1993) suggested that as few as 5 is usually sufficient. However, others have claimed many more are required (e.g Lesaffre & Spiessens, 2001). Table 1 presents results derived using only 3 quadrature points per latent variable. In contrast to the above arguments, when the models were estimated with increasing numbers of quadrature points (up to 20) there was little evidence of a convergence towards a stable set of results. Moreover, using adaptive quadrature (Liu & Pierce, 1994), which attempts to locate more of the points under the peak of the integrand, did not help.

We hypothesized that the party political trait latent variables, instead of being normally distributed, could be multimodal, with separate peaks for each party. We therefore explored models that made less restrictive assumptions about the shape of the latent political party trait distributions.

Latent class models

Models were specified with discretely distributed latent variables. With sufficient classes these correspond to non-parametric maximum likelihood (NPML; e.g. Laird, 1978) representations of a distribution of arbitrary form. The NPML estimate of the latent variable is based on a finite number of freely estimated integration points *c* (*c* = 1, …, C) with probabilities π_c and locations $\mathbf{z}_c$ so that the likelihood is equal to

$$\sum_c \pi_c \prod_j g\left(y_j \mid \mathbf{x}_j, \mathbf{u}_j = \mathbf{z}_c\right),$$

where *g*() is the multinomial logit link function and y_j is the conditional expectation of the response for respondent *j* given the sets of explanatory variables **x** and the shared random effects **u**.

The weighted locations of the points were constrained to a mean of zero on each latent dimension, so only C-1 location parameters per latent variable needed to be estimated. The variances and covariance of the latent variables were then derived from the locations of the quadrature points in the latent space. The probabilities of the points were constrained to sum to one, requiring only C-1 probability parameters.

The optimum number of integration classes was selected by using the Gateaux derivative method (Davis & Pickles, 1987; Heckman & Singer, 1984). Taking Model 2 as the base model two points were used initially. (Using just one point would have equated to a standard MNL model without random effects, shown earlier to be inadequate.) A further point with very low probability was then added and, keeping the fixed effects estimates constant, moved across a grid of locations (-10 to 10 in 40 equal steps per latent dimension) to see if any produced an improvement in likelihood. The location with the largest improvement, if any, was then taken as the starting value for a new integration point and the model re-estimated. This process was iterated until no further improvement in model fit could be achieved.

Estimating five classes did improve model fit over the four-class solution, but did not reach stable convergence, the final class having a very low probability and extreme location. For pragmatic reasons the four class solution was therefore accepted, even though it was not strictly the maximum likelihood estimate. This four-class model had a log-likelihood of -4545.45, a huge improvement over the likelihood of -4792.31 for Model 2 fitted with normally-distributed latent variables. However, the normally-distributed and discrete models are not nested and so straightforward likelihood ratio tests were not appropriate. To gauge relative model fit in this case we used the Bayesian Information Criterion (BIC; Schwarz, 1978). With N = 1914 and 11 estimated parameters Model 2 had a BIC of 9861.31. With the same N and 17 parameters Model 2d (d = 'discrete') had a BIC of 9219.37, affirming the superiority of the latent class model.

The discrete distribution of quadrature points can be interpreted as representing a number of latent classes that were homogeneous in their unobserved political party traits (though an interpretation as a parsimonious description of a potentially multimodal continuous distribution is perhaps more accurate). Figure 3 shows the locations and probabilities of the latent classes. The x-axis represents Labour- Conservative trait variation, the y-axis Liberal Democrat – Conservative trait variation.[2] Respondents were assigned to one of

[2] To allow the congruent positioning of the Labour latent class on the left and the Conservative latent class on the right of the figure the class locations were reflected about the zero point on each axis, so relative

the four classes by their maximum a posteriori probability of class entry based upon empirical Bayes estimation (e.g. Laird, 1982; Rabe-Hesketh, Pickles & Skrondal, 2003). The classes were then labeled according to their locations in the latent space and the modal party or parties of first preference.

---Figure 3 about here---

The Labour and Conservative classes were the largest, accounting for 41% and 31% of the respondents, respectively. Over 90% of the participants assigned to these classes were consistent in their party preference (when data were available for both years) and reported voting for their latent-class party. The third largest class, Liberal Democrat, had only 65% of those observed in both years consistently reporting the Liberal Democrats as their top-ranked party. Adding those respondents who only contributed to one wave of the data saw Liberal Democrat support rise to 74% in this class. 84% of votes in this class were for the Liberal Democrats, with virtually all of the remainder going to Labour.

The preference structure for the smallest class was more complex. Again, just considering those observed over both waves (which comprised 79% of the class total) the modal preference was for Labour in one wave and the Conservatives in the other (41%), split almost equally between shifts in both directions. The next most frequent preference was between the Conservatives and the Liberal Democrats (33%), again with switching direction split almost equally. Overall, 54% of votes in this class went to the Conservatives, with the remainder going 21% to Labour and 25% to the Liberal Democrats respectively. Vote was observed to almost always accompany preference.

Labour and Liberal Democrat preferences are negative and Conservative are positive. This was purely for display purposes; the arbitrary choice of Conservative in the MMNL model meant that relative

This class therefore appeared to contain respondents who had wholeheartedly switched preference and vote to/from the Conservatives, in roughly equal measure with both of the other parties. This class was labeled “Switchers”.

----Table 2 about here----

Table 2 shows the full model estimates from Model 2d in the second column. In comparison to Model 2 the Liberal Democrat latent trait variance in Model 2d was noticeably lower, and the correlation between latent dimensions was higher than that suggested by the model with normal latent trait distributions. Looking at the fixed effects log odds-ratio estimates, both Labour and the Liberal Democrats are strongly preferred overall to the Conservatives in 1997. However, the vote contrasts for both parties were negative, indicating that in comparison to the Conservatives neither party was garnering votes commensurate with their overall preferences. The preferences were broadly similar in 2001 but this time the vote contrast for the Liberal Democrats was not significant, i.e. compared to the Conservatives they were receiving about as many votes as their overall popularity would suggest.

Model 2d indicated that around the time of the 1997 General Election, when the Conservative party that had been in power for 18 years was defeated by a Labour landslide, Labour and Liberal Democrat supporters reported voting for their preferred party relatively less than did Conservative supporters. This could have reflected disillusionment on the part of some Conservatives supporters who, while expressing dissatisfaction with their party (and therefore relative satisfaction with the others), nonetheless went on to vote Conservative. Significantly for us, it could also reflect the

Conservative preferences are actually negative in the models.

influence of strategic voting on the part of Labour and Liberal Democrat supporters who voted, not in favor of their preferred party, but to oust the Conservatives.

Incorporating strategic voting

Alternative approaches to estimating the effects of strategic voting were discussed earlier, each of which can be incorporated here. The discrepancy between preference and vote (e.g. Black, 1978) is inherent in the models already presented and is indicated by the VOTE contrasts. The subjective approach classifies a voting decision as strategic if the respondent indicates this as their motivation for casting their vote (e.g. Evans and Heath, 1993). A dummy variable was therefore defined for those indicating a strategic motivation for their vote (STRAT=1, or 0 otherwise as determined by the Evans and Heath measure). Strategic motivation was thought likely to affect the vote decision itself rather than the underlying political preferences, so STRAT was specified to act only on the voting response (i.e. VOTE*STRAT). The effect of STRAT was not hypothesized to change over time, but it was hypothesized that the Liberal Democratic party would be subject to strategic voting to a greater degree than the other parties (see Spafford, 1972; Kim and Fording, 2001; Russell and Fieldhouse, 2002). Therefore separate coefficients for STRAT were estimated for each alternative, resulting in the fixed part of the model

$$Model2d{:}1 \quad V_{ij}^{a} = Model2d\ V_{ij}^{a} + x_{VOTE*STRAT\ j}\, g_{5}^{a}$$

With the random part identical to Model 2d, this gave Model 2d:1.

The third approach was to examine the impact of the objective conditions for strategic voting. This involved codifying this context in each electoral district and then including this information into the model, to see the extent to which it could account for the discrepancies between preferences and votes (c.f. Alvarez and Nagler, 2000). The concept of distance from contention (DFC; Cain, 1978; Niemi et al, 1992) was used to define the election context. DFC represents the relative strength of the parties in each electoral district based upon their previous election performance (here, the 1992 election for the 1997 wave and the 1997 election for the 2001 wave). DFC is computed by subtracting the number of votes polled by the second-placed party from the number of votes polled by each party in turn. So, if a party came first then DFC would be positive, if the party came second itself then DFC would be zero, and if the party came third it would be negative.

Because the MMNL models used here compares each party to a reference alternative, the difference in DFC scores between each party compared to the reference was needed. The relative DFC between two parties, (party1 – 2nd place) – (party2 – 2nd place), simplifies to (party1 – party2), i.e. the simple distance in votes between the parties or Prior Vote Margin (PVM). With the Conservatives as the reference party PVM was therefore positive if the party polled ahead of the Conservatives in the previous election and negative if behind. PVM was divided by 10,000 to put it in a scale comparable with the vote/preference measures. The mean values of PVM for Labour and the Liberal Democrats were 0.0914 and -1.0716 respectively, i.e. per constituency, Labour polled around 1000 votes more, and the Liberal Democrats about 10,000 votes less, than the Conservatives, on average.

PVM on its own did not capture the full electoral context because it did not encode the party positions; parties in two separate districts might have the same PVM but one might be in second place and the other third, say. According to many accounts of strategic voting within rational voter theory (e.g.McKelvey and Ordeshook, 1972; Cain, 1978) being in third place (or worse) is one of the most potent incentives for strategic voting. For this reason a dummy variable THIRD was defined, equal to 1 if the party was in third place in its electoral district and 0 otherwise. PVM, THIRD and their interaction were therefore election- and alternative-specific variables, as before loading only on the vote response. It was again hypothesized that the smaller Liberal Democrat party would prove more sensitive to strategic conditions than the larger parties. To allow the coefficients of these alternative-specific predictors to vary by party two dummy variables *lab* and *LD* were defined, 1 if the alternative was the Labour or Liberal Democrat parties, respectively, and 0 otherwise. This gave a further model, denoted Model 2d:2, with the fixed part (for Labour versus Conservative as an example)

$$Model2d{:}2 \quad V_{ij}^{lab} = Model2d\ V_{ij}^{lab} + x_{VOTE*PVM}^{lab} b_{1lab} + x_{VOTE*THIRD}^{lab} b_{2lab} + x_{VOTE*PVM*THIRD}^{lab} b_{3lab}$$

A final model was estimated, Model 2d:3, that combined Models 2d:1 and 2d:2 to give a model with all of the strategic variables STRAT, PVM, THIRD and PVMxTHIRD, to assess the relative effects of self-reported vs. contextual measures of strategic voting.

Table 2 shows the full results for Models 2d:1-2d:3 in columns 3-5. The addition of the self-described strategic voting dummy variable STRAT gave Model 2d:1 an improved fit over Model 2d, BICs of 9178.44 and 9219.37 respectively. The pattern of fixed effects was virtually the same over both models, with Labour and the Liberal Democrats being preferred overall to the Conservatives in both election waves but having lower vote contrasts (i.e. poorer conversion). The effect of STRAT was positive for the Liberal Democrats and negative for Labour, although significant only in the latter case. This suggested a net vote transfer away from Labour by self-reported tactical voters. This interpretation is supported by the rise in Labour's latent class probability from .41 to .49; it seemed that some strategically-voting Labour supporters were misclassified in Model 2d, which did not have the benefit of strategic voting information. The Liberal Democrat latent class probability dropped to .14 (from .22) indicating that a number of Liberal Democrat voters were also misclassified. There was also a slight reduction in the Conservative latent class probability from .31 to .28 indicating that some Conservative supporters voted tactically for the Liberal Democrats (see also Russell and Fieldhouse, 2004). The results from Model 2d:1 are consistent with the pattern of strategic voting that would be expected based upon the hypothesized positions of the parties in ideological space, providing evidence in support of the subjective measure of strategic voting.

Instead of self-reported strategic motivation, Model 2d:2 included the effects of PVM and THIRD, and so was based upon contextual rather than individual information. This model provided a huge improvement in fit compared to both the naïve Model 2d (without any strategic voting information) and Model 2d:1. For both Labour and the Liberal Democrats the effect of PVM was significant and positive, i.e. the probability of voting for either party increased with their vote from the previous election, in districts where the parties were in first or second place at the prior election. The coefficient for the Liberal Democrats was significantly higher than that for Labour. It therefore seemed, as expected, that the Liberal Democrat vote was much more sensitive to local party strength than was Labour vote (see MacAllister et al, 2001).

The effect of being THIRD in the prior election did not significantly harm the voting probability for the Labour party. For the Liberal Democrats, however, being in third place significantly reduced the voting probability, as would be expected. The effect of PVMxTHIRD, i.e. the effect of PVM when a party is in third place, was significant and positive on the Labour vote. Because PVM is always negative for parties in third place this implied a reduction in vote probability for Labour the greater this prior vote "distance" – i.e. sensitivity to the size of the vote deficit for Labour when in third place. The effect of PVMxTHIRD was non-significant for the Liberal Democrats, however, implying that when the Liberal Democrats are in third place the distance they are from the other parties does not make any additional difference for voting probability This is consistent with the notion of Duvergian strategic voting.

Unlike STRAT, taking into account these contextual effects altered the pattern of coefficients for the overall preferences and votes quite radically. As before, both the Liberal Democrats and Labour were significantly preferred to the Conservatives in both election years. However, the vote contrast coefficients, which were negative in Models 2d and 2s:1, were significant and positive for the Liberal Democrats. This suggested that after taking into account the relative strength of the parties at the prior election the Liberal Democrats were actually *more* likely to receive votes, given their overall level of preference, than were the Conservatives (see also Fieldhouse et al, 2004).

Model 2d:3 included STRAT, PVM and THIRD and again gave further improvement in model fit (BICs of 9059.92 vs. 9068.41 for Model 2d:2). This suggested that the individual-level and contextual-level variables were providing somewhat independent sources of information on the discrepancy between preference and vote. The pattern of significant fixed effects in this model was similar to that found in Model 2d:2. For the Labour effects the PVM coefficient was almost identical but no longer significant because of a larger standard error. The Liberal Democrat vote contrasts were still positive but only significant at the .10 level. STRAT was no longer significant for either party, nor was THIRD for the Liberal Democrats. The apparent dilution of the effect of these strategic variables when estimated together in the same model suggests that they share information, i.e. strategic voting occurs when the contextual factors militate for it.

Finally, Model 2d:3 was fitted again but this time including demographic and socio-economic covariates (measured in 1997, dummy variables for: sex, age (three variables), social class (two variables), tertiary education and homeownership). This final model was referred to as Model 2d:4. The covariates were included into the model as predictors of the latent classes, having their effects on the preference and voting responses indirectly via the latent political party traits.

Model 2d:4 again provided a significant improvement in log-likelihood (-4349.75) over the equivalent without covariates (Model 2d:3; -4435.50). However, because Model 2d:4 had nearly double the number of parameters (49) it had a worse BIC statistic, 9069.79 compared to 9059.92 for Model 2d:3.

Because the addition of the covariates did not improve model parsimony, for brevity only a summary of the results will be given here[3]. The locations and probabilities of the latent classes were much as found previously (changes of the order of 1% in class probabilities), although the estimated correlation between the latent traits increased to .96. The effects of

[3] Full results are available on request from the authors

the covariates were as would be expected: being younger, working class and a non-homeowner increased the probability of being in the Labour and Liberal Democrat latent classes compared to being in the Conservative class. Having a tertiary qualification increased the probability of Liberal Democrat class membership. Sex had no effect on class membership whatsoever. None of the covariates significantly predicted Switcher class membership over Conservative class membership, indicating the similarity of people allocated to these classes. Overall, the results from the final model confirmed the stability of the latent classes and also the similarity of the pro- and anti-Conservatives in terms of demographic and socio-economic information.

Estimating strategic vote

Figure 4 shows the relationship in the raw data between strategic motivations for vote and whether the vote was discrepant to the stated preference ranking, i.e. a non-preference vote.

--- Figure 4 about here

Taking the percentage of self-described strategic votes at face-value suggested that approximately 10% of votes were cast strategically. However, only 6% of votes were cast both against preference and declared to be strategic, whereas 15% were one or the other. Moreover, 4% of votes were declared strategic but cast according to stated preference. The majority (80%) of these ‘with-preference strategic voters’ selected an alternative party when subsequently asked what their “really preferred party” was, however.

Almost half of the non-preference votes (5% of total votes) were not declared strategic; when asked their motivation for their vote about 70% of these voters endorsed the statement “I voted for the best party.” This appears highly inconsistent; it may have represented measurement error in the party approval rankings or more likely in the motivation for vote question (Niemi et al, 1993), given the ambiguity of the word “best” in the above response (‘sincerely’ best or ‘strategically’ best, for instance).

The lower and upper limits on strategic voting suggested by these raw data lie between 6% (strategic *and* non-preference voters) and 15% (strategic *or* non-preference voters), in keeping with previous research (e.g. Heath & Evans, 1994; Clarke et al, 2004). It is clear that these raw data contain too much inconsistency and ambiguity to provide a more accurate estimate of strategic voting on their own (see also Blais & Nadeau, 1996). A more integrative approach was used by Alvarez and Nagler (2000) to provide model-based estimates of strategic voting. They looked how predictions of vote based upon their model of voting were influenced by removing strategically important covariates such as vote share.

Here, we took the parameter estimates from the most parsimonious model, Model 2d:3, as a starting point. Empirical Bayes scoring was used to give each individual an a posteriori vote probability for each party. The party with the highest probability was considered as that individual’s predicted vote. A second predicted vote was also made, but this time based upon a restricted model with the STRAT, PVM, THIRD and PVM*THIRD covariates set to zero, i.e. with the influence of the strategic variables removed. The two predicted votes were then crosstabulated to show the number of votes that were predicted to be “cast” for different parties when incorporating strategic information or not. (An important precondition for this approach was that the underlying model of observed vote be plausible; Model 2d:3 predicted 91.1% votes correctly which was considered acceptable.)

---- Table 3 about here

The crosstabulation of full versus restricted model-predicted votes is shown in Table 3. The entries off the main diagonal were the ones classified differently when taking into account the ‘strategic’ variables or not.

Of the predicted votes influenced by the strategic variables, the majority (191) was forecast for the Conservatives under the full model and for the Liberal Democrats under the restricted model. This was surprising as we would have expected that sensitivity to the strategic context would be manifest mostly between the Liberal Democrats and Labour. Further inspection revealed that each one of these 191 votes was cast by a voter allocated to the Switcher latent class. This class consisted of individuals who had preferred and voted for the Conservatives at least once across the two election years, but who had preferred and voted for another party in the other year. This led to their latent class being located somewhat centrally in latent space, at a point defined by the weighted average of the class members’ preferences and votes over the two election years. This point was closer to the location of the Conservative latent class than to the Labour latent class because of the preponderance of Conservative preferences and votes of its members. The limited influence of the latent trait for the Switcher class thus allowed the other covariates to exert a relatively greater influence than they did on the other latent classes, tipping the predicted vote more easily in the Conservative rather than the Labour direction.

On a more general level, the interpretation of the latent trait for the switcher class is one of latent instability of preference rather than just of latent party political preference; the election waves could have been modelled separately, but then the Switcher latent class would not have emerged, as the class members have consistent preferences and votes within individual elections. A very promising approach, beyond the scope of this paper, would be to explicitly model the changing structure in the latent classes over time (c.f. Vermunt, 2003).

Overall, slightly more than 9%[4] of all votes were classified differently by the two models, which is a plausible estimate of the influence of strategic factors on vote. As hypothesized, Liberal Democrat vote was the most sensitive to the strategic conditions; they were predicted to have 22% of the vote by the restricted model but only 18% with the full model (the lower figure mirroring their actual vote share over the two elections). This implies that the Liberal Democrats might have expected their vote share to have been 4% higher in the absence of strategic contextual factors and motivations, i.e. on a strategically ‘level’ playing field.

Conclusions

MMNL was found to be superior to MNL in modelling multiparty choice. Further, the assumption of smoothly varying, normally-distributed latent party preferences was also

[4] This figure was virtually the same, 8.9%, when estimated using model M2d:4, the model with demographic covariates

found to be untenable with these data. NPMLE allowed this assumption to be dropped and provided a huge improvement in model fit[5].

Based on the latent class models the Labour and Liberal Democrat classes were found to be distinct members of a predominately center-left bloc, with distinct patterns of preference but whose members seemed prepared to share votes strategically against the Conservatives. Although the Liberal Democrat class was found to exist between the two major parties in preference-space (and "nearer" to Labour than the Conservatives) it was found to be an oversimplification to place the parties on a single, left-right political continuum. Rather, the significant minority who did not have "single-peaked" preferences meant that models with two (albeit highly correlated) dimensions of latent political traits were required. This multidimensionality was attenuated with the addition of strategic-voting and demographic information into the models, however.

A naïve model of preference and vote that did not take the strategic local situation into account was found to be mis-specified. Information regarding strategic context was found to be more informative of the discrepancy between preference and vote than was a self-reported measure of strategic voting. This could mean that the majority of non-preference voters did not consider themselves to be strategic voters. However, this would suggest that the same factors that are assumed to influence strategic voting also strongly affect other types of non-preference voting. Although the subjective measure produced results that were difficult to interpret and that may have mis-estimated the prevalence of strategic

[5] It is important to note, however, that the evidence presented here does not allow us to argue for the "reality" of these discrete latent classes, but the approach does represent a useful and parsimonious way of representing the structure of political preferences. The names given to the classes are considered useful and plausible labels rather than as definitive identities.

voting it did improve model fit over and above the contribution of the purely contextual measures. However, these contextual measures, particularly the straightforward vote distance between the parties (PVM), although based on four- year-old data provided the greatest impact on model fit.

Comparing predicted votes based upon the full model and a model with no 'strategic' variables suggested that 9% of the votes were sensitive to these variables. This provided an estimate of strategic voting that combines non-preference voting, self-reported motivation as well as the objective contextual conditions. The use of latent variables for the underlying party preferences allowed these multiple, sometimes conflicting, sources of data to be used within a principled modeling framework. The explicit modelling of the changes in latent preference structure over time is a highly desirable next step for this approach.

As we hypothesized, consistent with the Duvergian model of strategic voting, it was the vote of the smallest of the three main parties that was found to be most sensitive to strategic factors, and the net losers of strategic voting.

References

Abramson, P.R., J.H. Aldrich, P. Paolino and D. Rohde. 1992. 'Sophisticated' Voting in the 1988 Presidential Primaries. *American Political Science Review* 86: 55-70.

Alvarez, R.M., and J. Nagler. 1995. Economics, issues and the Perot candidacy: voter choice in the 1992 presidential election. *American Journal of Political Science* 39:714–744.

Alvarez, R. M., and J. Nagler. 2000. A new approach for modelling strategic voting in multiparty elections. *British Journal of Political Science* 30:57-75.

Alvarez, R. M., J. Nagler, and S. Bowler. 2000. Issues, Economics and the Dynamics of Multiparty Elections: The British 1987 General Election. *American Political Science Review* 94:131–149.

Bhat, C.R. 2001. Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model. *Transportation Research B* 35 (7):677-693.

Black, J. H. 1978. The multicandidate calculus of voting. *American Journal of Political Science* 22 (3):609-638.

Blais, A., and R. Nadeau. 1996. Measuring Strategic Voting: A Two-Step Procedure. *Electoral Studies* 15:39-52.

Brynin, M. and D. Sanders. 1997. Party identification, political preferences and material conditions: evidence from the British Household Panel Survey, 1991-2. *Party Politics* 3: 53-77.

Cain, B. E. 1978. Strategic Voting in Britain *American Journal of Political Science* 22: 639-55.

Campbell, A., P.E. Converse, W.E. Miller, and D. Stokes. 1964. *The American Voter*. Chicago, IL: Chicago University Press.

Clarke, H.D., D. Sanders, M.C. Stewart, and P. Whiteley. 2004. *Political Choice in Britain*. Oxford: Oxford University Press.

Davis, R., and A. R. Pickles. 1987. A joint trip-timing store type choice model for grocery shopping, including inventory effects and nonparametric control for ommitted variables. *Transportation Research A* 21A (4/5):345-361.

Dow, J. K. and J. W. Endersby. 2004. Multinomial probit and multinomial logit: a comparison of choice models for voting research. *Electoral Studies* 23:107-122.

Duverger, M. 1954. *Political Parties: Their Organization and Activity in the Modern State*. New York, NY: Wiley.

Evans, G. and A. Heath. 1993. A Tactical Error in the Analysis of Tactical Voting: A Response to Niemi, Whitten and Franklin. *British Journal of Political Science* 23: 131-137.

Fieldhouse, E., A. Pickles, N. Shryane, J. Johnson, and K. Purdam. 2004. Party preferences and vote in a multi-party, simple plurality system: an analysis of Liberal Democrat support and vote using the generalized linear, latent and mixed models framework. *Political Studies Association Annual Conference 2004*, Lincoln. April 6th 2004.

Fisher, S. D. 2004. Definition and measurement of tactical voting: the role of rational choice. *British Journal of Political Science* 34 (1): 152-167.

Green, D., B. Palmquist, and E. Schickler. 2004. *Partisan Hearts and Minds*. New Haven, CT: Yale University Press.

Heath, A. , R. Jowell, J. Curtice, J. Field, and S. Witherspoon. 1991. *Understanding Political Change: The British Voter 1964-1987*. Oxford: Pergamon Press.

Heckman, J., and B. Singer. 1984. A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica* 52 (2): 271-320.

Hensher, D. A., and W.H. Green. 2003. The Mixed Logit model: The state of practice. *Transportation* 30: 133–176.

Johnston, R. J. and C. J. Pattie. 1991. Tactical Voting in Great Britain in 1983 and 1987: An Alternative Approach. *British Journal of Political Science* 21: 95-128.

Kendall, M. G. 1938. A new measure of rank correlations. *Biometrika* 30: 81.

Kim, H., and R. C. Fording. 2001. Does tactical voting matter? The political impact of tactical voting in recent British elections. *Comparative Political Studies* 34 (3): 294-311.

Laird, N. M. 1982. Empirical Bayes estimates using the non-parametric maximum likelihood estimate for the prior. *The journal of statistical computation and simulation* 15: 11-220.

Lesaffre, E., and B Spiessens. 2001. On the effect of the number of quadrature points in a logistic random-effects model: an example. *Journal of the Royal Statistical Society, part C (Applied Statistics)* 50 (3): 325-335.

Liu, Q., and D.A. Pierce. 1994. A note on Gauss-Hermite quadrature. *Biometrika* 81: 624-629.

Longford, N.T. 1993. *Random coefficient models*. Oxford: Oxford University Press.

Luce, R.D. 1959. *Individual Choice Behavior*. New York, NY: Wiley.

MacAllister I, E. Fieldhouse, A. Russell. 2002. Yellow fever? The political geography of liberal voting in Great Britain. *Political Geography* 21: 421-447.

McKelvey R.D. and P.C. Ordeshook. 1972. A general Theory if the Calculus of Voting in J.F. Hedon and L.J. Bernd (eds) *Mathematical Applications in Political Science VI*. Charlottesville, VA: The University of Virginia Press: pp. 32-78.

McFadden, D. 1974. Conditional logit analysis of qualitative choice behavior. In *Frontiers of econometrics*, edited by P. Zarembka. New York: Academic Press.

McFadden, D. and K. Train. 2000. Mixed Multinomial Logit models of discrete response. *Journal of Applied Econometrics* 15: 447–470.

Niemi, R.G., G. Whitten and M. N. Franklin. 1992. Constituency Characteristics, Individual Characteristics and Tactical Voting in the 1987 British General Election. *British Journal of Political Science* 22: 229-254.

Niemi, R. G., G. Whitten and M. N. Franklin. 1993. People in Glass Houses: A Response to Evans and Heath's Critique of our Note on Tactical Voting. *British Journal of Political Science* 23: 549-563.

Rabe-Hesketh, S., A Pickles, and A. Skrondal. 2001. GLLAMM: A class of models and a Stata program. *Multilevel Modelling Newsletter* 13 (1): 17-23.

Rabe-Hesketh, S., A. Pickles, and A. Skrondal. 2003. Correcting for covariate measurement error in logistic regression using non-parametric likelihood estimation. *Statistical Modelling* 3: 215-232.

Riker, W.H. 1982. The Two-party System and Duverger's Law: An Essay on the History of Political Science. *American Political Science Review* 76 (December): 753-766.

Russell, A., E. Fieldhouse, and I. MacAllister. 2002. The Anatomy of Liberal Support in Britain, 1974-97. *British Journal of Politics and International Relations* 4 (1): 49-74.

Russell, A., and E. Fieldhouse. 2004. *Neither left nor Right: The Liberal Democrats and the Electorate*. Manchester: Manchester University Press.

Sanders, D. 1996. New Labour, New Machiavelli - a cynic's guide to economic policy. *Political Quarterly* 67: 290-302.

Schulman, R.S. 1979. A geometric model of rank correlation. *American Statistician* 33: 77-80.

Schwarz. 1978. Estimating the dimension of a model. *Annals of statistics* 6: 461-464.

Skrondal, A. and S. Rabe-Hesketh. 2003. Multilevel logistic regression for polytomous data and rankings. *Psychometrika* 68 (2): 267-287.

Spafford, D. 1972. Electoral Systems and voter's behaviour. *Comparative Politics* 5: pp.129-134.

StataCorp. 2003. *Stata Statistical Software: Release 8.0*. College Station, TX: Stata Corporation.

Stroud, A.H. 1971. *Approximate Calculation of Multiple Integrals*. Englewood Cliffs, NJ: Prentice- Hall.

Thomson, K., A. Park, and L. Brook. 1999. British General Election Study 1997: Cross-section Survey, Scottish Election Study, and Ethnic Minority Election Study. London: National Centre for Social Research.

Vermunt, J.K. 2003. Multilevel Latent Class Models. *Sociological Methodology*, 33, 213-239.

Whitten, G., and H. Palmer. 1996. Heightening Comparativists' Concern for Model Choice: Voting Behavior in Great Britain and the Netherlands. *American Journal of Political Science* 43: 1231-1247.

Zhang, J. 2004. Binary choice, subset choice, random utility, and ranking: A unified perspective using the permutahedron. *Journal of Mathematical Psychology* 48 (2): 107-134.

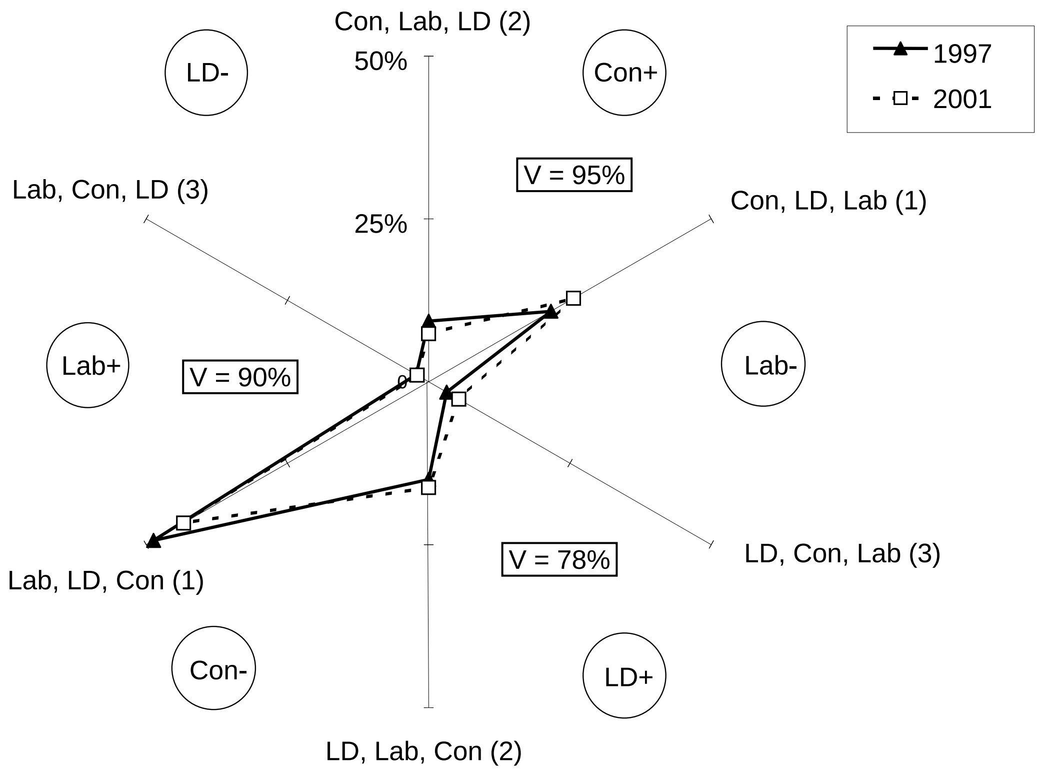

Figure 1: Variation on a permutahedron showing percentages of party preference rank-orders for BEPS-2 waves 1997 and 2001, and percentage of vote for most preferred party (V)

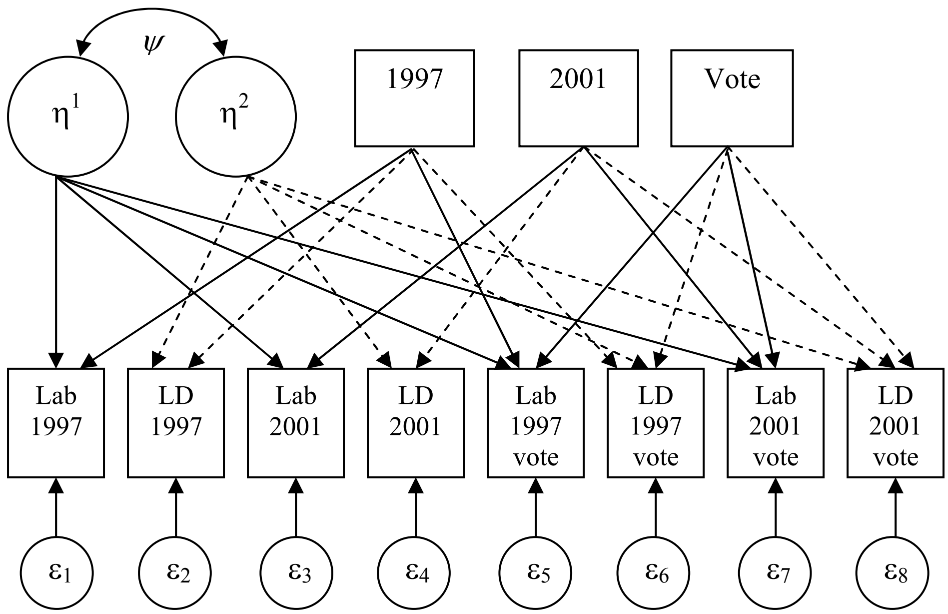

Figure 2: Model 2. Mixed Multinomial Logit model of ranked preference and vote for the Labour (Lab) and Liberal Democratic (LD) parties (in relation to the Conservative party) in 1997 and 2001. Arrows denoting Lab effects are solid, LD arrows are dashed. Larger circles represent latent party political traits for Lab (η_1) and LD (η_2) (and their correlation, ψ); smaller circles represent error in the fixed effects estimates (ε_1-ε_8).

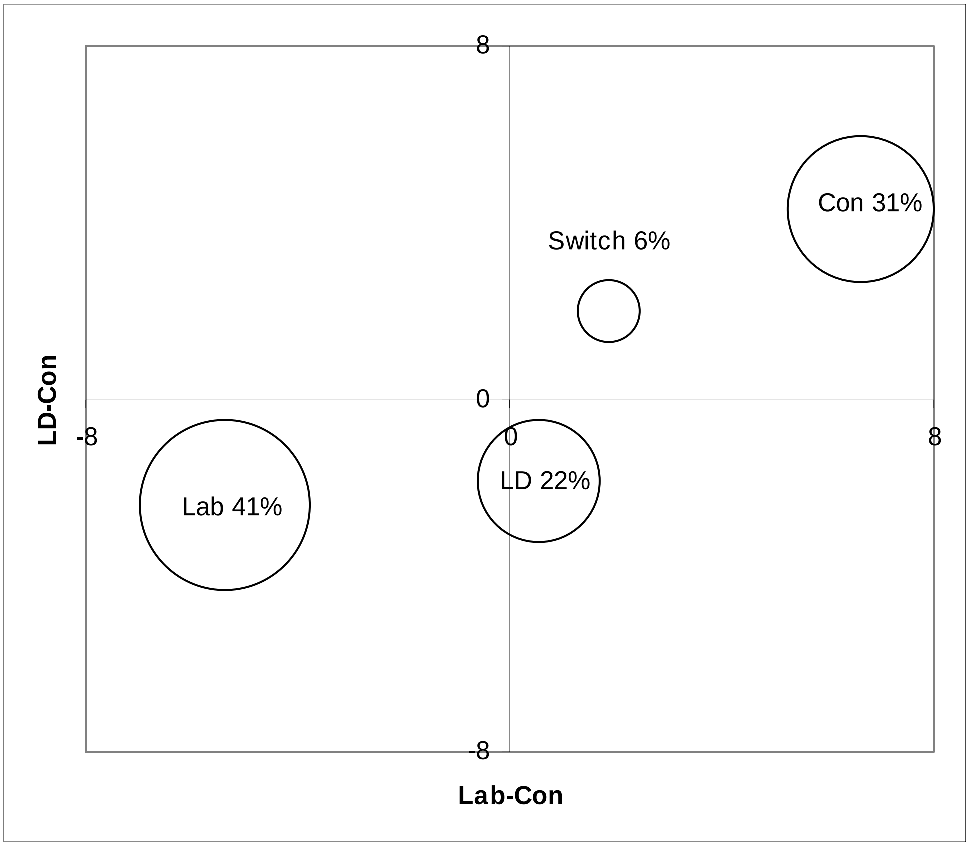

Figure 3: Graph of latent political party trait space from Model 2d. The x-axis is the Lab-Con dimension, the y-axis the LD-Con dimension. The four latent classes are plotted in this space, labeled with plausible party / ideological affiliation (Lab = Labour, LD = Liberal Democrat, Switch = Party Switchers, Con = Conservative) and the percentage of respondents in the class.

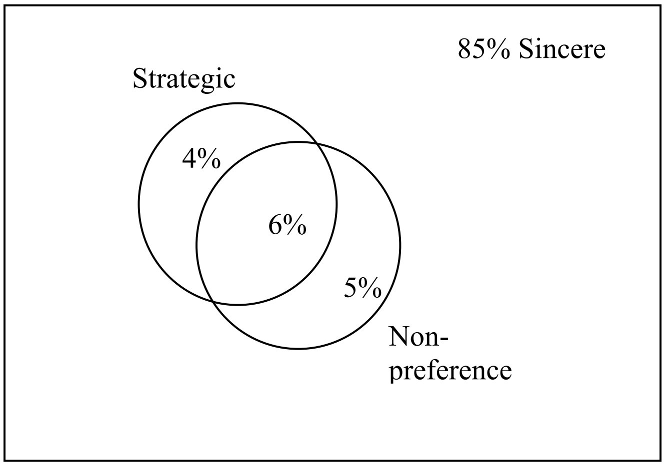

Figure 4: Reported vote % classified by self-reported strategic motivation and whether the vote was cast against stated party preference (i.e. non-preference). N = 2892 voting occasions, figures rounded to nearest %

Table 1: Model fit and random effects estimates for models of preference and vote

	Model 2	Model 1	Model 0
Parameters	11	10	8
Log-likelihood	-4889.09	-5202.82	-7434.93
η^1 variance	25.33	17.73	-
(SE)	(1.51)	(0.89)	
η^2 variance	14.14		-
(SE)	(1.10)		
λ^2 factor loading	-	0.66	-
(SE)		(0.01)	
$\psi_{\eta^1 \eta^2}$ correlation	.81	-	-

Notes. N = 1914 voters and N = 2892 voting occasions for all models.

Table 2: Estimates for Models with discrete latent variable distribution

Effect	Model 2d		Model 2d:1		Model 2d:2		Model 2d:3	
	Est.	SE	Est.	SE	Est.	SE	Est.	SE
Fixed Effects								
Labour (vs. Con.)								
1997	2.45**	.26	1.69**	.29	1.81**	.33	1.72**	.34
1997 vote contrast	-0.70**	.15	-0.55**	.16	-0.06	.18	-0.02	.20
2001	2.27**	.29	1.44**	.29	1.52**	.34	1.45**	.35
2001 vote contrast	-0.35**	.15	-0.11	.17	-0.20	.20	-0.13	.23
STRAT	-	-	-1.53**	.56	-	-	-1.17	.74
PVM[1]	-	-	-	-	0.27**	.11	0.26	.17
THIRD	-	-	-	-	0.15	.75	0.51	.60
PVMxTHIRD[1]	-	-	-	-	0.71**	.32	0.70**	.26
LibDem (vs. Con.)								
1997	1.23**	.17	1.05**	.25	0.82**	.27	0.95**	.32
1997 vote contrast	-0.54**	.16	-0.80**	.17	1.03**	.28	0.68*	.38
2001	1.40**	.22	1.20**	.26	0.92**	.29	1.07**	.36
2001 vote contrast	-0.18	.17	-0.36**	.16	0.99**	.24	0.67*	.37
STRAT	-	-	0.68	.51	-	-	0.56	.59
PVM[1]	-	-	-	-	0.98**	.18	0.87**	.15
THIRD	-	-	-	-	-1.06**	.38	-0.76	.61
PVMxTHIRD[1]	-	-	-	-	-0.40	.30	-0.24	.37
Random Effects								
Trait Labour (η^1)	25.42		23.51		27.55		24.75	
Trait LD (η^2)	9.02		12.99		13.51		12.93	
Correlation $\psi_{\eta^1\eta^2}$	.92		.91		.93		.93	
Latent class probability[2]								
Labour	.41		.49		.45		.48	
Liberal Democrat	.22		.14		.18		.15	
Switchers	.06		.09		.09		.09	
Conservative	.31		.28		.28		.28	
Parameters	17		19		23		25	
Log Likelihood	-4545.45		-4517.43		-4447.30		-4435.50	
BIC	9219.37		9178.44		9068.41		9059.92	

Notes. N = 1914 voters and N = 2892 voting occasions for all models.
Significant effects: * p < .10, ** p < .05
[1] Effect of prior vote margin is per 10,000 votes.
[2] Maximum a posteriori probability based on Empirical Bayes estimation.

Table 3: Predicted vote based upon full and restricted* Model 2d:3

Full model predicted vote	Restricted Model* predicted vote			
	Con	Lab	LD	Total (%)
Con	845	0	191	1,036 (36)
Lab	0	1,340	5	1,345 (46)
LD	0	69	442	511 (18)
Total (%)	845 (29)	1409 (49)	638 (22)	2,892 (100)

* Vote predicted from parameters estimated for Model 2d:3, but with STRAT, PVM, THIRD and PVM*THIRD covariates set to zero. 9.2% off-diagonal votes, i.e. those predictions affected by ‘strategic’ covariates

Supplemental material – available on request

Table Xa: Estimates for Model with demographic covariates (Model 2d:4)

Effect	Lab-Con		LD-Con	
	Est.	SE	Est.	SE
Fixed Effects				
1997	-0.46	.45	-0.69	.38
1997 vote contrast	-0.03	.21	**0.69**	**.30**
2001	-0.73	.45	-0.56	.39
2001 vote contrast	-0.13	.26	**0.67**	**.29**
STRAT	**-1.02**	**.45**	0.61	.35
PVM	0.25	.13	**0.86**	**.16**
THIRD	0.47	.59	-0.72	.42
PVMxTHIRD	**0.70**	**.28**	-0.20	.29
Random Effects				
Latent trait η^a	26.25		14.17	
Correlation $\psi_{\eta^1\eta^2}$	.96			
Parameters	49			
Log Likelihood	4349.75			
BIC	9069.79			

Notes. N = 1914 voters and N = 2892 voting occasions for all models. Significant effects (p < .05) are shown in **bold**. [1] Effect of prior vote margin is per 10,000 votes.

Supplemental material – available on request

Table Xb: Covariate effects for Model 2d:4

	Latent Class							
	Lab		LD		Switcher		Con	
	Est	SE	Est	SE	Est	SE	Est	SE
Latent class probability[2]	.48		.15		.09		.28	
Lab-Con (η^1) location	6.74	-	2.47	-	0.03	-	-5.09	-
Ld-Con (η^2) location	4.03	-	4.84	-	0.33	-	-3.88	-
Covariates								
FEMALE	-0.07	.12	0.03	.16	-0.07	.22	-	-
AGE 18-30	**0.57**	**.21**	**0.67**	**.28**	0.55	.34	-	-
AGE 31-45	**0.79**	**.16**	**0.71**	**.22**	0.18	.30	-	-
AGE 46-60	**0.36**	**.16**	0.38	.23	0.07	.29	-	-
AGE > 61	-	-	-	-	-	-	-	-
GHClass Salariat	-	-	-	-	-	-	-	-
GHClass Intermediate	0.17	.15	0.15	.21	-0.06	.27	-	-
GHClass Working	**1.16**	**.19**	**0.70**	**.26**	0.43	.34	-	-
TERQUAL	0.21	.16	**0.79**	**.21**	0.48	.28	-	-
NOHOMEOWN	**0.80**	**.16**	**0.44**	**.21**	0.16	.30	-	-
Class intercept	**-0.40**	**.18**	**-1.51**	**.27**	**-1.18**	**.31**	-	-

Notes. N = 1914 voters and N = 2892 voting occasions for all models. Significant effects ($p < .05$) are shown in **bold**. [2] Maximum a posteriori probability based on Empirical Bayes estimation.