Multilevel Modelling of Political Party Preference

Nicola Marie Crane

2011

Contents

Abstract **3**

1 Introduction **4**

1.1 General Theories of Political Party Preference 4

1.2 Individual-Level Predictors of Party Preference 5

1.3 Regional Differences in Party Preference 7

1.4 Overall Project Aims 8

2 Data **9**

2.1 The British Social Attitudes Survey, 2005 9

2.2 The Present Analyses 10

3 Analytic Methodology **13**

3.1 Binary and Multinomial Logistic Regression 13

3.2 Multilevel Models 14

3.3 Goodness of Fit Statistics 15

3.4 Computer Software 16

4 Results **18**

4.1 Exploratory Analyses 18

4.2 Regression Models 18

5 Discussion **28**

References **32**

6 Appendices **34**

6.1 Appendix A - Explanatory Variables 34

Abstract

Predicting political party preference has been a topic of interest to many researchers, but the scope of analysis has widened in recent years with the development of software capable of handling complex multilevel analyses where data is hierarchically structured. Previous research has found a range of factors consistently useful in predicting party preference, with social class being one of the most widely-cited influences. However, even when numerous variables are accounted for, regional differences in party preference have frequently been found. The analyses presented here aimed to explore the nature of social class related voting, and also to examine regional differences in political party preference. The analyses were conducted using a combination of logistic regression models, and the multilevel extensions of these models. It was found that class-related voting was significant when measured using individuals' self-reported social class, but not their earnings, indicating that its roots may lie in social identity rather than socioeconomic status. Homogeneity within regions was not found for all political parties in the binary models, and a multinomial model showed no effect of region. These findings indicate that the predictors of politically party preference are not universal to all parties, despite arguments otherwise in the majority of previous research. Regional differences only affect some parties, and this should be examined in more depth in future research.

Chapter 1

Introduction

1.1 General Theories of Political Party Preference

Numerous theories of voting behaviour have been proposed over the years, and a brief summary of the predominant theories is presented below.

Before the 1940s, it was believed that individuals' political party preferences were based upon straightforward choices between parties, determined by information acquired through parties' advertising campaigns. However, this theory fell out of favour when it was found that people's party affiliations remained fairly stable, unaffected by such influences (Antunes, 2010). One group of theories that emerged to explain this are known as the sociological models, which focuses on characteristics such as socioeconomic status and religion to explain voting choice. Such theories suggest that people with similar characteristics tend to show homogenous voting patterns within that group (Lazarsfeld, Berelson, & Gaudet, 1944). One thing, however, that sociological theories of voter choice do not account for is that although certain demographic characteristics can be used to predict the homogeneity of such groups' preferences, the specific parties that those groups vote for change over time.

An alternative to the sociological models of voter behaviour are the socio-psychological models, developed shortly after the sociological model became popular. These models take the focus away from group characteristics and place a lot more emphasis on individual opinions and beliefs. These theories argue that people choose their party affiliation because they agree with the policies of the parties and simply choose the one closest to their own opinions. These

beliefs and opinions, however, are formed on the basis of group membership and socialisation (Miller & Shanks, 1996). However, these theories differ from the sociological models as they argue that political party choice is more of a conscious choice than the former school of thought would imply. Individuals act in a way which is consistent with that of a group which they believe they belong to, and part of this regulation of behaviour includes party affiliation (Antunes, 2010).

Irrespective of the underlying reasons, these theories and the results of empirical research show that individuals with certain characteristics in common tend to display similar patterns of voting behaviour. The next section explores some of the specific characteristics which have been found to influence political party affiliation. It should be noted that the terms 'party affiliation', 'party preference', 'voter intent' and 'voting behaviour' are used interchangeably here. Although it is common for opinion polls of voting intention to reflect different outcomes to actual voting behaviour in elections, previous research (Rose & McAllister, 1990) has found little difference in the variables that influence individuals' voting intention and actual voting behaviour.

1.2 Individual-Level Predictors of Party Preference

Social Class

As mentioned above, research has found that individuals from a similar social class background tend to show similar patterns of political party preference (Lazarsfeld et al., 1944). More recent research has examined this relationship in more detail, and Jones, Johnston, and Pattie (1992) highlight that it has been observed frequently that there is a positive relationship between the probability of voting Labour and being working class, and a similar relationship regarding middle class individuals voting Conservative.

This trend of 'class voting' has been argued by some to be in decline since around the late 1960s (Franklin & Mughan, 1978). A popular theory to explain this trend is known as 'class dealignment thesis' (Clark & Lipset, 1991). This argument posits that social class is no longer important in modern society and that distinct differences between members of different social classes have diminished and have far less impact upon individual political

affiliation than they used to. However, Heath, Evans, and Payne (1995) argue that these changes are simply "trendless fluctuations" and there has been no significant change, and Manza, Hout, and Brooks (1995) go as far as to say that "in no democratic capitalist country has vote been entirely independent of class in a national election."

The concept of social class is problematic, as it is commonly operationalised by examining the individual's occupation. Franklin and Mughan (1978) highlight the fact that using occupation as a proxy for social class is the normal convention when modelling voter intention. However, Samar, Azimi, and Dadvand (2007) demonstrate that individuals' subjective perceptions of their own and others social class are affected by a number of individual characteristics not simply restricted to occupation. It seems logical, therefore, that individuals' self-reported social class is a useful variable to examine when modelling voter preference with the caveat that this distinction should be noted when making comparisons with other studies of the relationship between social class and political party choice.

Other variables

Although social class is the most commonly cited predictor of political party preference, a number of other variables have been found to be useful in determining the parties that individuals are most likely to vote for. For instance, age is an important factor, and Jones et al. (1992) found that older voters were more likely vote Conservative than younger voters. Additionally, gender has been linked to party choice, with early studies finding that women had stronger right wing preferences than men, with this relationship being reversed in more recent years. However, Norris (1985) argues that this earlier trend was found because women have a longer lifespan than men and these studies had not taken this into account. Additionally, the gender gap has been shown to be larger in America, with European women more likely to vote for left-wing parties than men, but less so than American women.

Rose and McAllister (1990) emphasise the importance of family on voting behaviour and argue that parental political allegiance is a predictor of individual party choice, although there is not a great deal of research examining this link. Religious beliefs have been used to predict political party choice in the past, with more religious people being found to be more likely to vote Conservative (Lewis-Beck, 1986), although it has been argued that the

link between religion and politics is more complicated than such a straightforward direct relationship (Kotler-Berkowitz, 2001).

The link between educational level and political allegiance is a fairly complex one. Whitten and Palmer (1996) highlight that although in other countries it has been found that the more educated people are, the more likely they are to vote for non-mainstream parties or those with explicitly secular values; in the UK it has been found that more years in formal education has a positive relationship with the likelihood of voting Conservative.

1.3 Regional Differences in Party Preference

Jones et al. (1992) argue that when modelling the patterns in voting behaviour, it is essential that both individual-level variables and higher-level variables, such as region or country, are taken into account. They argue that the effects of different regions on political party choice have not been accounted for by individual-level effects. The aforementioned research concerned voting behaviour in the UK, but similar patterns have emerged in America (Morrill, Knopp, & Brown, 2007), and again they persist even when various demographic features of different regions are taken into account. Johnston, Pattie, and Allsopp (1988) argues that the wealth of evidence for regional effects suggests that a multi-level modelling approach is needed in order to get an accurate picture of voter characteristics in the UK, and this will be discussed later.

This tendency for heterogeneity in the way different variables affect voting behaviour in different regions has been attributed to the "neighbourhood effect" (Johnston, 2005) which simply proposes that individuals are likely to show similar patterns of voting behaviour to those around them. Agnew (1996) takes a stronger position and argues that locality is intrinsically linked to political behaviour and highlights evidence from research conducted in Italy where different areas with similar demographic characteristics of their populations had radically different party affiliations. The reasons for these differences have been widely debated, with some suggesting that they are caused by "localised opinion voters" whereas others claim that it is more to do with "party-identity voters", those who have deeper links to particular parties. Regardless of which theory best explains the phenomena, evidence still indicates regional differences unaccounted for by other factors.

1.4 Overall Project Aims

This project aims to examine the variables which influence political party affiliation. In order to compare the results with previous theoretical and empirical work, individual-level variables in the analysis will basic demographic characteristics such as age, sex and income, as well as identity-related variables such as the individual's perception of their own social class membership, and other identity-defining groups. In addition, the effects of regional variation will be taken into account to assess to what extent it affects political party affiliation. The next chapter will discuss the data set used for analysis and elaborate more on the variables to be included in the models.

Chapter 2

Data

2.1 The British Social Attitudes Survey, 2005

The data analysed came from the 2005 British Social Attitudes Survey (BSAS). This survey is conducted annually in the UK since 1983 by the National Centre for Social Research (NCSR), an independent research institute. Each year, over 3000 people from England, Scotland and Wales are surveyed using a combination of face-to-face interview, and a self-completed questionnaire. The survey contains questions regarding opinions on social and political issues as well as individuals' characteristics such as age, sex and occupation. There were four different versions of the 2005 survey and each contained the same basic demographic questions, but different questions about individuals' views on social and political issues.

The 2005 version of the survey was chosen for a number of reasons. Firstly, this was the most recent year that the survey results were available for in which there had been a general election. Using the survey from an election year means that people were more likely to have been engaged in politics at the time, and our findings can be used to make inferences about the actual election results. In 2005, the BSAS was conducted between June and September, and as the general election in the UK in 2005 was on May 5th, the survey closely coincides with this date.

Data collection

Respondents were selected randomly using the Postcode Address File, a list of all postal addresses in the United Kingdom. Although buildings in the sample were selected at random, individual respondents were not. Where there were multiple flats or dwellings within one building, interviewers used a Kish grid to randomly select which flat or dwelling to interview an individual from. Where there were multiple individuals within the same household, the same procedure was used to select an interviewee.

The face-to-face component of the interview took place in the interviewee's home (or on their doorstep) and responses were recorded using a Computer Assisted Personal Interview program on a laptop. Upon completion of the interview, respondents were given a questionnaire to complete and return at a later date. There were four different versions of the questionnaire which were randomly assigned to respondents.

2.2 The Present Analyses

Data cleaning and recoding

The original sample consisted of responses from 4268 individuals. However, the dataset used in the analyses contained data from 1944 individuals, as some cases had to be removed. As mentioned before, several version of the questionnaire had been administered and respondents who had not been asked all of the questions represented by the variables of interest were excluded.

A number of cases had to be removed as the responses were not suitable for analysis. For example, there was no logical way to recode responses from those who refused to give their age or gave a 'don't know' response to certain questions and so all cases with these responses were removed from the dataset. Other explanatory variables were also recoded or some cases being removed; these will be discussed further below. To summarise the changes, all unordered nominal variables were recoded into dummy variables. Most ordinal variables were left unchanged, although cases with 'don't know' or 'other' responses which do not fit into ordinal ranking were removed. The variable representing educational attainment was reversed so that the higher the level the education, the higher the associated numerical value;

this was done for ease of analysis.

Explanatory Variables

The dataset contained over 1000 different variables which had been recorded. The variables included in the analyses were based upon those related to those used as predictors of political party preference in previous research. Explanatory variables chosen were self-reported social class, father's party preference, mother's party preference, marital status, economic activity, sex, age, highest educational achievement, earnings in the last year, religion, religiosity, and a number of identity categories. The complete list of explanatory variables and their levels, can be found in Appendix A.

Region

The region variable had 11 different options; all of the 9 standard Government Office regions for England, plus Scotland and Wales. These regions are: Scotland, Wales, North West England, North East England, Yorkshire and the Humber, East Midlands, West Midlands, East, London, South East and South West. These were the smallest sized regional divisions of the country available in the dataset.

Response Variable

The response variable for this analysis is the political party that the respondent most closely identified with. This was originally coded with the first category, which would automatically become the reference category, as 'Conservative', but was recoded so that 'Labour' would be the reference category. This was due to the fact that Labour was the most common party choice and so it is logical, for ease of analysis, that we examine why people chose alternatives to the most popular choice.

Aims

Previous research concerning political party choice in the UK has usually focussed on the factors that influence people to vote for either of the two main political parties - Labour and

the Conservatives. However, in recent years, the Liberal Democrat party has increased in popularity, after they merged with the Social Democratic Party (not to be confused with a party of the same name which formed after the merger) in 1988. Since then, they managed to increase their share of the vote from 17.8% and 20 seats in the House of Commons in 1992, to 23% of votes and 57 seats in 2010.

Much of the research concerning political party preference focuses on the left-wing/right-wing dichotomy that Labour and the Conservatives loosely fit into, and very little research examines voting influences for more than two parties. Additionally, although much theoretical work acknowledges the crucial role that regional differences have to play in political party preference, multilevel modelling approaches still have not been widely used in order to investigate these differences more in-depth.

As stated earlier, the aim of this project is to examine the variables that are linked with political party affiliation. This will be done via a number of methods. Firstly, one-level and two-level binomial logistic regression models will be fitted to the data. These models will allow us to examine the variables which are associated with party affiliation for each of the different parties. Conducting both single-level and multi-level analyses will allow us to compare the models, and examine whether regional differences play in important role for all parties.

Secondly, one-level and two-level multinomial logistic regression models will be fitted to the data. These models will allow us to explore the reasons that individuals voted for specific parties as constrasted with the most popular choice, Labour. Again, the two-level model will allow us to examine regional differences, and can be compared to the single-level model to examine whether these differences have a significant impact on party choice.

Chapter 3

Analytic Methodology

3.1 Binary and Multinomial Logistic Regression

Analyses of political party choice and voter intent are often carried out using binary logistic regression. It is part of a family of generalized linear models which extend the basic general linear model by relaxing the assumption that the response variable must be normally distributed, and allows for the inclusion of nonlinear link functions. Binary logistic models can be expressed in the following form:

$$logit(p_i) = log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 x_{1i} + ... + \beta_n x_{ni} + \epsilon_i$$

where $logit(p_i)$ is the log odds of success or in this case, a 'yes' response, calculated from the odds ratio (i.e. $\frac{p_i}{1-p_i}$) of that response being in either (i.e. yes or no) category of response. The log odds are linked to the regression covariates $x_1....x_n$ via a logit transformation. In the above notation, i represents an individual response, n represents the number of covariates included, β_0 represents the intercept, $\beta_1...\beta_n$ the coefficients of the covariates, and ϵ_i the error term.

Although this allows us to examine political party choice by comparing those who vote for a particular party with those who do not vote for that party, it does not allow us to make comparisons between those who vote for a party and those for vote for a specific alternative. In other words, although we can, for example, look at the differences between those who vote Labour and those who do not, binomial logistic regression models are restrictive in that they do not allow us to compare between Labour voters and Conservative voters. Dow and

Endersby (2004) argue that binary analyses of voter behaviour often portray the data in terms of the main party and all of their opposition, and homogenises separate groups with distinct characteristics which should not be ignored.

Multinomial logistic regression models allow us to model data with a polytomous response variable whilst making comparisons between one category of response, the reference category, and the other response categories. This model can be expressed in the following form:

$$log\left(\frac{\pi_i^{(s)}}{\pi_i^{(t)}}\right) = \beta_0^{(s)} + \beta_1^{(s)} x_{1i} + ...\beta_n^{(s)} x_{ni} + \epsilon_i^{(s)}, t = 1, s = 2, ..., t$$

where $\frac{\pi_i^{(s)}}{\pi_i^{(t)}}$ is the odds of a 'yes' response for category s of the response variable in proportion to the odds of a 'yes' response for category t of the response variable, which has been used as a reference category.

An alternative to the multinomial logistic regression model is the multinomial probit model, which differs in the link function used. Although it has been argued that it is better for modelling voter behaviour (Quinn, Martin, & Whitford, 1999), Dow and Endersby (2004) argue that the inherent relative simplicity of the logit model and numerous issues that arise when trying to apply the probit model to this type of data, mean that the logistic link model is the superior choice.

3.2 Multilevel Models

Multi-level models can be used to extend linear and non-linear models so that variation at more than one level can be incorporated. They are commonly used in educational research to assess the amount of variance that can be accounted for by differences between classes and schools. A typical model would include individual pupils within classes and classes within school. With the development of appropriate software, the use of multilevel techniques to examine voting behaviour has increased in popularity. Jones et al. (1992) argue that there are important regional differences in predictors of party choice that cannot be identified unless we use multilevel models.

The notation for the level-one multilevel binary logistic regression model is expressed thus:

$$log\left(\frac{p_{ij}}{1-p_{ij}}\right) = \beta_{0j} + \beta_{1j}x_{1ij} + ... + \beta_{nj}x_{nij} + \epsilon_{ij}$$

with the level two model being:

$\beta_{0j} = \gamma 00 + U_{0j}$

$\beta_{1j} = \gamma 10$

The level one model is similar to the single-level binary logistic regression mode, except for that now the individual response, i is nested within the region, j. Additionally, each region now has its own specific error term, ϵ_{ij}. The intercept, β_{0j} is also region-specific and its components can be seen in the level 2 model. It is compromised of $\gamma 00$, the level 2 intercept, also known as the grand mean, and U_{0j}, the level 2 error term.

The multilevel multinomial model can be expressed as:

$$log\left(\frac{\pi_{ij}^{(s)}}{\pi_{ij}^{(t)}}\right) = \beta_{0j}^{(s)} + \beta_{1j}^{(s)}x_{1ij} + ...\beta_{nj}^{(s)}x_{nij} + \epsilon_{ij}^{(s)}, t = 1, s = 2, ..., t$$

which is similar to the basic multinomial logistic model, but here i is again the individual voter, j is the region, and the random individual error term ϵ_{ij} can again be different for each level. In the equation, as before, s is between 2 and t (as category $t = 1$ of the response variable is used as a reference category). The level two model is:

$\beta_{0j}^{(s)} = \gamma 00^{(s)} + U_{0j}^{(s)}$

$\beta_{1j}^{(s)} = \gamma 10^{(s)}$

This is identical to that of the multilevel binomial logistic regression model, except for the inclusion of s, the category being constrasted with the reference category.

3.3 Goodness of Fit Statistics

Residual Intraclass Correlation Coefficients

The residual intraclass correlation coefficient is used to assess the amount of variance attributable to variance between the level 2 clusters (in a 2-level model), in this case, the different regions. As the multilevel models that will be fitted to the data all assume a logistic distribution for the individual level (level 1) variables, the residual intraclass correlation coefficient is expressed as:

$$\rho = \frac{\sigma_{u0}^2}{\sigma_{u0}^2 + \pi^2/3}$$

where σ^2_{u0} represents the level 2 variance and $\sigma^2_{u0} + \pi^2/3$ represents the overall variance within the model. The intraclass correlation coefficient can be described as 'unconditional' when calculated from the null model, as no parameters have yet been added to the model, and 'conditional' when describing a model which contains explanatory variables.

Akaike Information Criterion and the Likelihood-ratio test

The Akaike Information Criterion (AIC) is used for comparing different models to assess which one best fits the data. The lower the AIC value, the better the fit. It is sometimes considered better than simply comparing model deviances as it takes the number of parameters into account and penalises models with too many parameters and so reduces the chance of an overfitted model being selected. The AIC value for each model is calculated thus:

$AIC = 2k - 2ln(L)$

with k representing the number of parameters in the model and $-2ln(L)$ being the deviance of the model. The AIC values will be examined when comparing one- and two-level models in order to assess which best describes the data. Although likelihood ratio tests will also be used, the AIC values will be viewed alongside these to examine if the same conclusions can be drawn when the extra parameter of the level-2 variable, region, is taken into account.

The likelihood ratio test can be expressed as:

$$D = -2ln\frac{\text{likelihood of null model}}{\text{likelihood of alternative model}}$$

and simply compared the log-likelihoods of the two models. It is assumed to approximate the chi-square distribution and so a p value can be generated, with degrees of freedom (df) being calculated from the df for the alternative model minus the df for the null model.

3.4 Computer Software

Single-level analyses were conducted in Stata, using the logistic or mlogit procedure.

All multilevel analyses were conducted using the gllamm procedure in Stata. gllamm stands for Generalized Linear Latent And Mixed Models and allows multilevel models to be fitted to data which has a hierarchical or multilevel structure, using maximum likelihood estimation. When dealing with complex models, it can be difficult or time consuming to

accurately generate maximum likelihood estimates, and so gllamm uses adaptive quadrature, and repeatedly tries a series of integrations until it converges on a single value.

Chapter 4

Results

4.1 Exploratory Analyses

Exploratory analyses were conducted to get an initial idea of the trends in the dataset for some of the explanatory variables. Cross tabulation was used to explore the differences in political party choice by region, and the frequencies and their associated percentages can be found in Table 4.1. There seems to be some regional differences in political party preferences. Figure 4.1 shows a bar chart of political party choice by sex and indicates very little differences in males and females. In Figure 4.2, we can see a bar chart of party preference by self-reported social class. There seems to be definite differences between the classes, with middle class people having a higher tendency to vote Conservative, but working class people more likely to vote Labour. Interestingly, those who gave a 'don't know' or 'not applicable' response to being asked what social class they considered themselves a member of were most likely to choose 'other' parties than Labour, Conservative or Liberal Democrat.

4.2 Regression Models

As the aim of this project was to examine the impact of regional difference on political party choice, the analyses focussed on whether or not there was significant variation between regions. In order to answer this question, firstly one- and two-level binary logistic regression models were fitted to the data. The AIC values of the two models were compared to identify which

Figure 4.1: Political Party by Sex

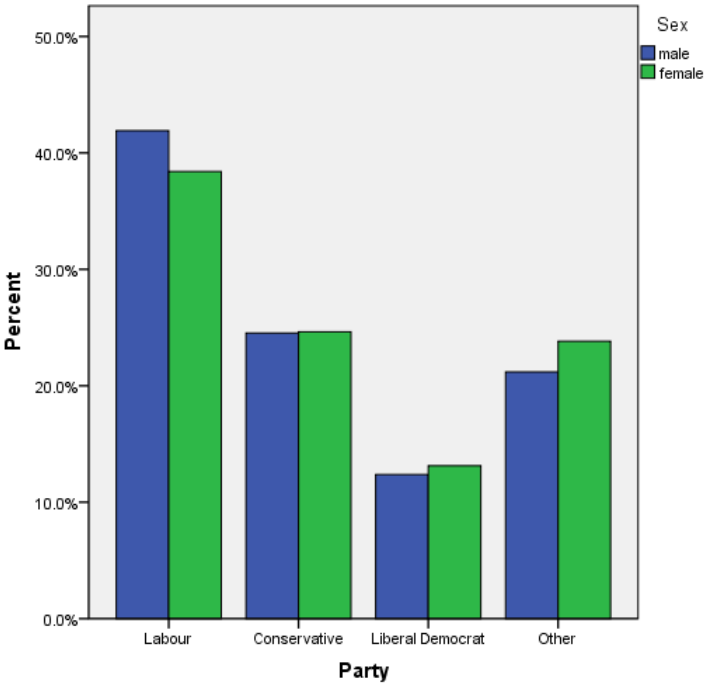

Figure 4.2: Political Party by Class

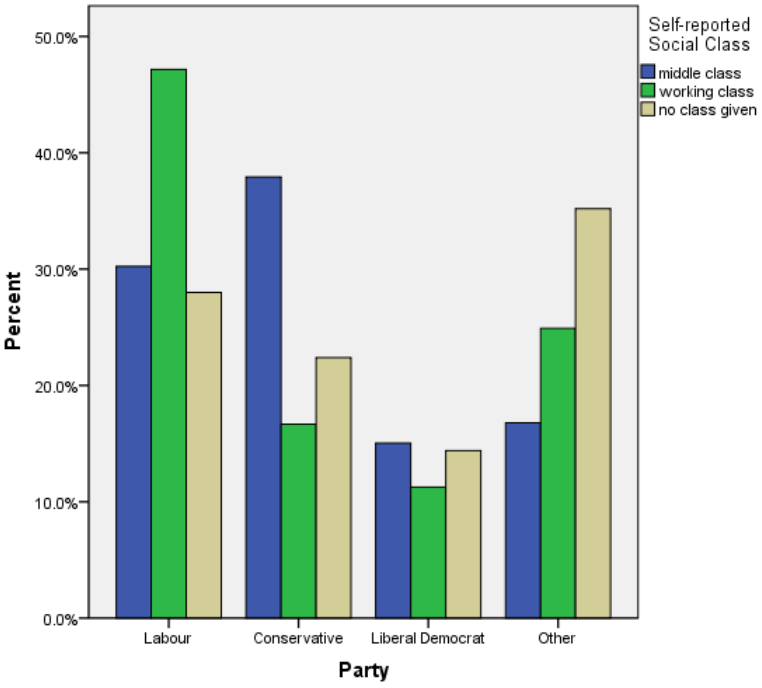

Table 4.1: Table of party choice by region

	Party				
	Labour	Conservative	Lib Dem	Other	Total
Scotland	72 (37.3%)	29 (15.0%)	17 (8.8%)	75 (38.9%)	193 (100%)
North East	76 (59.4%)	6 (4.7%)	11 (8.6%)	35 (27.3%)	128 (100%)
North West	99 (45.2%)	48 (21.9%)	31 (14.2%)	41 (18.7%)	219 (100%)
Yorkshire/Humber	81 (47.4%)	30 (17.5%)	21 (12.3%)	39 (22.8%)	171 (100%)
West Midlands	70 (41.9%)	50 (29.9%)	13 (7.8%)	34 (20.4%)	167 (100%)
East Midlands	56 (35.2%)	42 (26.4%)	23 (14.5%)	38 (23.9%)	159 (100%)
Eastern	34 (41.5%)	22 (26.8%)	9 (11.0%)	17 (20.7%)	82 (100%)
South West	47 (27.2%)	60 (34.7%)	38 (22.0%)	28 (16.2%)	173 (100%)
South East	124 (31.8%)	141 (36.2%)	49 (12.6%)	76 (19.5%)	390 (100%)
London	73 (44.2%)	35 (21.2%)	28 (17.0%)	29 (17.6%)	165 (100%)
Wales	44 (45.4%)	15 (15.5%)	9 (9.3%)	29 (29.9%)	97 (100%)
Total	776 (39.9%)	478 (24.6%)	249 (12.8%)	441 (22.7%)	1944 (100%)

model best fitted the data, and likelihood ratio tests were carried out to formally test these differences. Additionally, the final multilevel models were compared to their corresponding null multilevel models to assess the unconditional intraclass correlation before parameters were added to the model.

The multilevel analyses were carried out using the gllamm procedure in Stata. Due to the large number of variables and the time-consuming nature of the gllamm procedure, a forward-selection procedure was used. Initially, univariate analyses were carried out, with all of the variables modelled separately. Variables that were not significant at this point were eliminated from the analyses. Those variables that remained, starting with the most significant, were added to the model one at a time, again with only variables that were statistically significant at a 5% level being retained.

Binary logistic regression models

One- and two-level binary logistic regression models were fitted to the data in order to examine which variables predict affiliation to each political party, and whether these effects vary between different regions. Parameter estimates were fairly similar between the corresponding one- and two-level models and all parameters that were significant in one were significant in the other.

Table 4.2 shows the significant parameters and associated coefficients for the model

concerning a preference for the Conservative party. The results shows that individuals who are married are more likely to vote for the Conservative party, as are older people, Christian and Jewish people, those who describe themselves as middle class, those whose mother or father supported the Conservative party, and self-identified Eurosceptics and fox hunting supporters. However, anti-war campaigners and feminists were much less likely to vote for the Conservative party than for other parties.

Table 4.2: Binary logistic models of Conservative party affiliation

	One-level model			Two-level model		
Parameter	Coefficient	SE	p value	Coefficient	SE	p value
Intercept	−3.637	0.237	< 0.001	−3.737	0.279	< 0.001
Marital-Yes	0.354	0.125	0.005	0.316	0.128	0.013
Age	0.019	0.004	< 0.001	0.019	0.004	< 0.001
Religion-Christian	0.483	0.134	< 0.001	0.489	0.136	< 0.001
Religion-Jewish	1.399	0.697	0.045	1.428	0.710	0.044
Class-Middle	0.863	0.125	< 0.001	0.838	0.128	< 0.001
Father-Conservative	0.856	0.164	< 0.001	0.870	0.167	< 0.001
Mother-Conservative	0.716	0.167	< 0.001	0.673	0.169	< 0.001
ID-Antiwar campaigner	−0.924	0.267	< 0.001	−0.933	0.267	< 0.001
ID-Eurosceptic	0.866	0.158	< 0.001	0.890	0.161	< 0.001
ID-Feminist	−0.849	0.318	0.008	−0.919	0.325	0.005
ID-Fox hunt supporter	1.239	0.187	< 0.001	1.218	0.190	< 0.001

Table 4.3 shows the variables linked to Labour party support. Individuals who have been divorced, consider themselves to be working class, and whose mothers were Labour supporters were more likely to vote for Labour. However, individuals whose fathers voted Conservative and considered themselves Eurosceptics or fox hunting supporters were less likely to vote for Labour.

Table 4.4 shows that there were far fewer parameters in the model predicting Liberal Democrat support than in the binary regression models for other parties. The results show that the more educated individuals were, the more likely they were to support the Liberal Democrats. Also, if their father supported the Liberal Democrats (or similar pre-existing parties) or they were an anti-war campaigner, this also increased the odds of them supporting the Liberal Democrats.

Table 4.3: Binary logistic models of Labour party affiliation

	One-level model			Two-level model		
Parameter	Coefficient	SE	p value	Coefficient	SE	p value
Intercept	−0.782	0.099	< 0.001	−0.760	0.108	< 0.001
Marital-Divorced	0.438	0.145	0.002	0.430	0.146	0.003
Class-Working	0.524	0.103	< 0.001	0.507	0.104	< 0.001
Father-Conservative	−0.640	0.136	< 0.001	−0.618	0.137	< 0.001
Mother-Labour	0.824	0.107	< 0.001	0.807	0.108	< 0.001
ID-Eurosceptic	−0.854	0.164	< 0.001	−0.858	0.165	< 0.001
ID-Fox hunt supporter	−0.831	0.208	< 0.001	−0.811	0.209	< 0.001

Table 4.4: Binary logistic models of Liberal Democrat party affiliation

	One-level model			Two-level model		
Parameter	Coefficient	SE	p value	Coefficient	SE	p value
Intercept	−2.622	0.127	< 0.001	−2.648	0.149	< 0.001
Highest educational qual	0.207	0.036	< 0.001	0.205	0.037	< 0.001
Father-Lib Dem/Lib/SDP	1.230	0.261	< 0.001	1.226	0.264	< 0.001
ID-Antiwar campaigner	0.886	0.201	< 0.001	0.912	0.202	< 0.001

Table 4.5 shows the variables associated with supporting parties other than Labour, the Conservatives or the Liberal Democrats. The results give us more of an indication of people unlikely to vote for these parties than those likely to vote for them. The older an individual, and the more educated they were, the less likely they were to vote for 'other' parties. Additionally, those who were Christians, in employment, considered themselves to be middle class, those whose fathers were Conservative supporters, or those whose mothers were Labour supporters were less likely to vote for 'other' parties.

In order to assess whether the addition of region as a level-2 variable significantly improved the fit of the model, single level binomial logistic regression models, which did not take region into account, were fitted to the data. In order to compare the models, the AIC values were compared and, in addition, likelihood ratio tests were carried out. The AIC values for the one- and two-levels models, and the p-values based upon chi-square contrasts from the likelihood ratio tests can be seen in Table 4.6.

These results show that the addition of region as a higher-level variable only significant

Table 4.5: Binary logistic models of other party affiliations

	One-level model			Two-level model		
Parameter	Coefficient	SE	p value	Coefficient	SE	p value
Intercept	1.218	0.233	< 0.001	1.232	0.241	< 0.001
Age	-0.025	0.004	< 0.001	-0.025	0.004	< 0.001
Highest educational qual	-0.171	0.037	< 0.001	-0.178	0.037	< 0.001
Religion-Christian	-0.465	0.118	< 0.001	-0.456	0.119	< 0.001
Class-Middle	-0.351	0.133	0.008	-0.322	0.135	0.017
Father-Conservative	-0.465	0.152	0.002	-0.446	0.154	0.004
Mother-Labour	-0.737	0.129	< 0.001	-0.784	0.132	< 0.001
Work-Employed	-0.310	0.127	0.015	-0.297	0.128	0.021

Table 4.6: AIC values for the 1- and 2-level binomial logistic regression models

Model	AIC (1 level)	AIC (2 level)	Likelihood ratio test p-value
Labour	2364.882	2363.820	0.080
Conservative	1718.526	1697.648	< 0.001
Liberal Democrat	1417.494	1415.82	0.055
Other parties	1940.471	1927.665	< 0.001

improved the fit of the models concerning Conservative and 'other' party affiliations. Both the likelihood ratio tests and a basic comparison of the AIC values support this, with the non-significant models showing little change in the AIC value for the two-level models when compared with the one-level models, but the significant models showing a decreased in AIC value.

Next, the conditional and unconditional intraclass correlation coefficients were calculated for each of the multilevel binomial logistic regression models to assess the level of homogeneity within regions before and after the relevant parameters for each model were considered. The results can be seen in Table 4.7 and they indicate that, Labour, Conservative and Liberal Democrat support does not vary much between regions with an extremely low unconditional ICC value, accounting for less than 0.1% of the variance in each model. However, the conditional ICC, based on the models with significant explanatory variables, is much higher for all of these parties. Regarding support for 'other' parties, the unconditional ICC is high, accounting for 21% of the variance in this model. However, when parameters are added to the model, this value drops to accounting for only 2.3% of the variance.

Table 4.7: Intraclass correlation coefficients for null and final multilevel binary logistic regression models

Model	ICC (Null)	ICC(Final)
Labour	0.001	0.011
Conservative	0.001	0.058
Liberal democrat	0.001	0.017
Other parties	0.210	0.023

When viewed in conjunction with Table 4.6, it can be seen that the only models which are significantly improved by including region as a level- 2 variable are the ones for Conservative support, for which region accounts for 5.8% of the variance and 'other' party support, for which region accounts for 2.3% of the variance. To conclude, Labour and Liberal Democrat support is best modelled using a standard logistic regression approach, whereas Conservative party and 'other' party support require a multilevel approach for the model to best fit the data.

Multinomial Logistic Regression Models

Although we have explored the reasons that individuals vote for specific parties, previous literature, as mentioned earlier, suggests that a multinomial modelling approach best captures the differences between the different parties. As in the previous section, both one- and two-level models were fitted to the data, and AIC and ICC values were calculated in order to examine which model best describes the data.

The parameters and associated estimates for the one-level model can be seen in Table 4.8. The equivalent multilevel multinomial model can be seen in Table 4.9. Statistically significant values are highlighted in grey.The AIC value for the one-level multinomial model was 4469.029, and for the two-level model was 4468.903, indicating that the two-level model does little to improve the fit. A likelihood ratio test leads us the same conclusion, $\chi^2(1, N = 1944) = 2.13, p = 0.145$.

Additionally, the ICCs for the null and final multilevel models were compared. The null model had an ICC of 0.005 and the final model had an ICC of 0.007. These results indicate that there was very little regional variance in the data, and the addition of parameters to the model did not account for or reveal any significant amount of regional variance. Therefore, a

one-level model is adequate to describe this data.

For the part of the model which describes which respondents voted for the Conservative party as opposed to Labour, marital status was in part significant, with individuals who were divorced or widowed less likely to vote Conservative. Those who described themselves as Eurosceptics or fox hunting supporters were more likely to vote Conservative, and environmentalists and anti-war campaigners and self-described members of the working class were less likely to vote Conservative. Those whose mothers or fathers voted Conservative were more likely to have a preference themselves for that party as opposed to Labour, and those who could not recall their mother having a preference for any political party were also more likely to vote Conservative than Labour.

Those who described themselves as Eurosceptics were more likely to vote for the Liberal Democrats than Labour, as were those whose parents had a preference for Liberal Democrats or similar parties. Interestingly, those whose mother had a preference for the Conservatives were more likely to vote for the Liberal Democrats than for Labour. Individuals who considered themselves to be working class were less likely to vote for the Liberal Democrats than for Labour.

Regarding the part of model comparing those who preferred the 'other', less mainstream parties to Labour, educational qualifications and age were negatively associated with 'other party' support, with older people and those with more education less likely to choose 'other' parties'. Christians were less likely to choose these parties rather than Labour, as were people who were divorced, middle class or working class (both here as the reference category was 'don't know'). Eurosceptics were more likely to prefer an 'other' party to Labour, as were all of those whose mother's voted for any party other than Labour or had no preference.

Table 4.8: One-Level Multinomial Model of Political Party Affiliation

	Conservative versus Labour			Lib Dem versus Labour			'Other' versus Labour		
Parameters	Coefficient	SE	p value	Coefficient	SE	p value	Coefficient	SE	p value
Intercept	-1.897	0.562	< 0.001	-0.931	0.618	0.132	1.135	0.452	0.012
Highest educational achievement	-0.105	0.129	0.415	-0.036	0.139	0.795	-0.500	0.116	< 0.001
Age	0.010	0.008	0.218	-0.011	0.010	0.250	-0.026	0.007	< 0.001
Religion-Christian	0.257	0.148	0.082	-0.227	0.159	0.153	-0.428	0.131	0.001
Marital-Divorced	-0.792	0.221	< 0.001	-0.318	0.225	0.157	-0.530	0.192	0.006
Marital-Widowed	-0.636	0.246	0.010	-0.382	0.310	0.218	-0.304	0.253	0.230
Identity-Eurosceptic	1.365	0.201	< 0.001	0.484	0.246	0.049	0.785	0.215	< 0.001
Identity-Fox hunting supporter	1.463	0.238	< 0.001	0.272	0.325	0.402	0.354	0.283	0.211
Identity-Environmentalist	-0.492	0.172	0.004	0.108	0.181	0.551	-0.229	0.166	0.168
Identity-Antiwar campaigner	-0.723	0.296	0.014	0.555	0.239	0.021	0.024	0.238	0.920
Class-Middle	0.167	0.307	0.587	-0.411	0.328	0.210	-0.585	0.271	0.031
Class-Working	-0.812	0.302	0.007	-0.684	0.317	0.031	-0.684	0.252	0.007
Father-Conservative	0.939	0.199	< 0.001	0.391	0.241	0.105	0.133	0.214	0.535
Father-Lib Dem/Liberals/SDP	-0.327	0.437	0.454	0.821	0.391	0.036	0.063	0.409	0.877
Mother-Conservative	1.299	0.228	< 0.001	0.668	0.264	0.011	0.981	0.232	< 0.001
Mother-Lib Dem/Liberals/SDP	0.773	0.374	0.039	0.836	0.378	0.027	0.748	0.378	0.048
Mother-Other	0.946	0.689	0.170	1.041	0.675	0.123	1.508	0.555	0.007
Mother-None/NA	0.695	0.171	< 0.001	0.327	0.185	0.077	0.991	0.145	< 0.001

Table 4.9: Multilevel Multinomial Model of Political Party Affiliation

	Conservative versus Labour			Lib Dem versus Labour			'Other' versus Labour		
Parameters	Coefficient	SE	p value	Coefficient	SE	p value	Coefficient	SE	p value
Intercept	−1.920	0.564	0.001	−0.956	0.621	0.123	1.127	0.454	0.013
Highest educational achievement	−0.968	0.130	0.457	−0.025	0.140	0.859	−0.495	0.116	< 0.001
Age	0.010	0.008	0.230	−0.012	0.010	0.242	−0.026	0.007	< 0.001
Religion-Christian	0.255	0.148	0.085	−0.229	0.160	0.151	−0.428	0.132	0.001
Marital-Divorced	−0.780	0.221	< 0.001	−0.313	0.225	0.165	−0.523	0.192	0.006
Marital-Widowed	−0.619	0.247	0.012	−0.360	0.310	0.246	−0.283	0.254	0.265
Identity-Eurosceptic	1.367	0.202	< 0.001	0.486	0.246	0.048	0.788	0.215	< 0.001
Identity-Fox hunting supporter	1.436	0.239	< 0.001	0.248	0.326	0.447	0.326	0.284	0.250
Identity-Environmentalist	−0.493	0.172	0.004	0.108	0.181	0.552	−0.225	0.166	0.175
Identity-Antiwar campaigner	−0.729	0.296	0.014	0.547	0.240	0.022	0.016	0.238	0.945
Class-Middle	0.174	0.308	0.573	−0.400	0.329	0.224	−0.579	0.272	0.033
Class-Working	−0.790	0.303	0.009	−0.662	0.318	0.037	−0.665	0.252	0.008
Father-Conservative	0.936	0.199	< 0.001	0.384	0.241	0.111	0.126	0.214	0.555
Father-Lib Dem/Liberals/SDP	−0.325	0.438	0.459	0.817	0.393	0.038	0.058	0.411	0.887
Mother-Conservative	1.262	0.228	< 0.001	0.633	0.265	0.017	0.947	0.232	< 0.001
Mother-Lib Dem/Liberals/SDP	0.761	0.374	0.042	0.824	0.379	0.030	0.740	0.378	0.051
Mother-Other	0.902	0.690	0.191	0.999	0.675	0.139	1.464	0.557	0.009
Mother-None/NA	0.677	0.171	< 0.001	0.312	0.186	0.093	0.977	0.145	< 0.001

Chapter 5

Discussion

The overall conclusion to be drawn from the results is that political party preference cannot simply be described in terms of a two-party dichotomy with the same variables predicting preference for all parties. Additionally, although regional differences have been widely acknowledged, it should not be assumed that they exist for all political parties or work in the same way.

The binary logistic regression models showed that regional differences occurred for some political parties, but not for all of them. The intraclass correlation coefficients indicated that no regional differences were visible for Labour, the Conservatives or the Liberal Democrats in the null model, despite exploratory analyses seeming to indicate otherwise. However, when the relevant parameters were added, differences emerged for the Conservative party, with almost 6% of the variance in the model accounted for by differences between different regions. These findings differ from those of previous research, as reviewed earlier, which indicated a general regional difference for all political parties. Additionally, there was a fairly large amount of regional homogeneity for 'other' parties, which decreased when parameters were entered into the model. This makes sense when we examine the nature of this response category, as 'other' encompasses a wide range of different parties, some of which only run for election or focus their campaigning in specific areas.

Some of the significant variables in the binary models were consistent with previous research, and others were not. Although being a Christian was positively associated with voting Conservative, this finding is problematic to interpret due to the way in which this

variable was coded. All denominations of Christianity were collapsed into one variable. However, previous research has shown that specific Christian denominations are affiliated with particular political parties, for example, Catholics are more likely to vote Conservative and Protestants are more likely to vote Labour (Kotler-Berkowitz, 2001). Interestingly, Judaism was also significant for Conservative affiliation, which was not discussed by any of the authors reviewed earlier.

Previous research (Antunes, 2010) argues that political party preference is intrinsically linked with an individual's sense of identity and this seems to be reflected in the results of the research presented here. All of the models for the three main parties include at least one identity category as a statistically significant predictor. In addition, self-reported social class was statistically significant for the binary models for both Labour and Conservative party preference. Not only does this fit in with a wealth of previous findings, it also provides some support for the idea that class voting is more to do with identity than socioeconomic status. Although a variable representing income was originally included in the model, it was found not to be statistically significant.

As also found in previous research (Rose & McAllister, 1990), parental political party affiliation had an effect on their children's party preference with the party affiliation of at least one parent being statistically significant in the corresponding model for each of models for the three main parties. Another aspect of these results that supports the theory that voting behaviour is influenced by social identity is the degree to which parental party affiliation affects individual's party preference. Not only are people likely to vote in the same way as their parents, as shown by the binary models, but the multinomial model indicates that parental affiliation with a party is linked with a lower probability of voting for a 'rival' party. This can be seen in the contrasts comparing Conservative preference with Labour preference, in which individuals whose parents voted for the Liberal Democrats were more likely to vote Conservative rather than Labour. Similarly, in the contrasts between Liberal Democrats and Labour, if an individual's mother voted Conservative, that individual was more likely to vote for the Liberal Democrats as opposed to Labour. This seems to indicate some degree of partisanship in party support. An alternative explanation is that the Liberal Democrats and the Conservatives lie closer to each other than to Labour on the political spectrum; however,

there still may be some influence of social identity in determining allegiance to these parties.

The multinomial model contrasted Labour, the predominant party with each of the other response options. Regional differences were not significant in the model, implying that people preferred particular parties to Labour for fairly consistent reasons in different areas.

One criticism of the sampling method is that using the Postcode Address File may result in an unrepresentative sample. Homeless people who have no registered address would not be eligible for inclusion, for example. However, it could be argued that this is irrelevant as they would therefore not be able to vote anyway. However, there are other inherent problems with the use of this database. Any building which contains multiple properties, (e.g. a block of flats, or student halls of residence) will only appear as the single building in this database, and only one flat or dwelling was sampled when multiple existed within the same building, and so people living in this kind of accommodation are much less likely to be included in the sample.

A potential issue with the interpretation of the results is the 'identity' variables. Although they seem to be useful in predicting party preference, and indicate a link between party preference and social identity, they may instead reflect individuals voting on social or political 'issues', rather than any genuine group membership.

A better-fitting model could perhaps have been acquired if more information was available about regional variables, such as levels of unemployment and population density. Lewis-Beck (1986) draws attention to past studies which have included these variables and argues that evidence shows that they have a significant effect on political party affiliation. Additionally, level 2 differences may have been larger if information was available for smaller areas, such as constituencies or towns. However, although these responses were originally recorded, for the purpose of maintaining respondent anonymity, they were excluded from the publicly available dataset.

In conclusion, these results have supported previous findings that regional differences do exist in political party preference, and remain even when a number of variables are accounted for. However, these regional differences do not apply to all parties, and care should be taken in the future to not imply that either regional differences or the factors predicting party preference are the same for all political parties. Future research should examine whether or

not smaller scale place homogeneity exists, for example, within towns rather than regions. In addition, social class and similar identity based categories seemed to influence party preference. However, the exact nature of this link is somewhat vague and could be attributed to other factors, and so a deeper theoretical understanding of this relationship could be acquired in future research by specifically asking respondents questions about why they voted for a particular party and what their second and third choices would be. This kind of analysis has been conducted previously using multinomial techniques (Whitten & Palmer, 1996), although did not include an in-depth analysis of the underlying reasons behind the ordered choices.

References

Agnew, J. (1996). Mapping politics: how context counts in electoral geography. *Political Geography*, *15*, 129–146.

Antunes, R. (2010). Theoretical models of voting behaviour. *Exedra*, *4*, 145–170.

Clark, T., & Lipset, S. (1991). Are social classes dying? *International Sociology*, *6*, 397–410.

Dow, J., & Endersby, J. (2004). Multinomial probit and multinomial logit: a comparison of choice models for voting research. *Electoral Studies*, *23*, 107–122.

Franklin, M., & Mughan, A. (1978). The decline of class voting in britain: Problems of analysis and interpretation. *The American Political Science Review*, *72*, 523–534.

Heath, A., Evans, G., & Payne, C. (1995). Modelling the class-party relationship in britain, 1964-92. *Journal of the Royal Statistical Society*, *158*, 563–574.

Johnston, R. (2005). Anglo-american electoral geography: Same roots and same goals, but different means and ends? *Professional Geographer*, *57*, 580–587.

Johnston, R., Pattie, C., & Allsopp, J. (1988). *A nation dividing?: the electoral map of great britain, 1979-1987*. Longman.

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

Kotler-Berkowitz, L. (2001). Religion and voting behaviour in great britain: A reassessment. *British Journal of Political Science*, *31*, 523–554.

Lazarsfeld, P. F., Berelson, B., & Gaudet, H. (1944). *The peoples choice: how the voter makes up his mind in a presidential campaign*. New York: Columbia University Press.

Lewis-Beck, M. (1986). Comparative economic voting: Britain, france, germany, italy. *American Journal of Political Science*, *30*, 315–346.

Manza, J., Hout, M., & Brooks, C. (1995). Class voting in capitalist democracies since world war ii: Dealignment, realignment or trendless fluctuation? *Annual Review of Sociology*, *21*, 137–162.

Miller, W. E., & Shanks, J. M. (1996). *The new american voter*. Cambridge, MA: Harvard University Press.

Morrill, R., Knopp, L., & Brown, M. (2007). Anomalies in red and blue: Exceptionalism in american electoral geography. *Political Geography*, *26*, 525–553.

Norris, P. (1985). The gender gap in britain and america. *Parliamentary Affairs*, *38*(2), 192.

Quinn, K., Martin, A., & Whitford, A. (1999). Voter choice in multi-party democracies: a test of competing theories and models. *American Journal of Political Science*, *43*, 1231-1247.

Rose, R., & McAllister, I. (1990). *The loyalties of voters: A lifetime learning model*. Sage London.

Samar, G., Azimi, H., & Dadvand, B. (2007). Socioeconomic status and class perception. *Iranian Journal of Language Studies*, *3*, 183–200.

Whitten, G., & Palmer, H. (1996). Heightening comparativists' concern for model choice: Voting behavior in great britain and the netherlands. *American Journal of Political Science*, *40*, 231–260.

Chapter 6

Appendices

6.1 Appendix A - Explanatory Variables

Table 6.1: Explanatory Variables

Variable	Levels
Self-reported social class	No class given
	Working class
	Middle class
Father's party preference	Labour
	Conservative
	Lib dem/Liberals/SDP/Alliance
	Other
	None/refused
Mother's party preference	Labour
	Conservative
	Lib dem/Liberals/SDP/Alliance
	Other
	None/refused
Marital status	Married/Living as married
	Separated/Divorced
	Widowed
	Never married
	Other/refused
Economic activity	Full time education/training

	In work/waiting to take up work
	Unemployed
	Retired
	Other/refused
Religion	Atheist
	Christian
	Jewish
	Hindu
	Muslim
	Other
	Don't know/refused
Sex	Male
	Female
Highest achieved educational qualification	None
	CSE or equivalent
	O Level or equivalent
	A Level or equivalent
	Higher education below degree
	First degree
	Postgraduate degree
Earnings in last year	< 3999
	4000 - 5999
	6000 7999
	8000 9999
	10000 11999
	12000 14999
	15000 - 17999
	18000 - 19999
	20000 - 22999
	23000 - 25999
	26000 - 27999
	29000 - 31999
	32000 - 37999
	38000 - 43999
	44000 - 49999
	50000 - 55999

	56000 or more
	Refused to answer
Religiosity	Very religious
	Somewhat religious
	Not very religious
	Not at all religious
Age in years	(Numeric value)
Identity	None of these
	Animal rights campaigner
	Anti-war campaigner
	Environmentalist
	Eurosceptic
	Feminist
	Supporter of fox hunting
	Graduate
	Vegetarian/vegan