Electoral participation of South Asian communities in England and Wales

Edward Fieldhouse and David Cutts

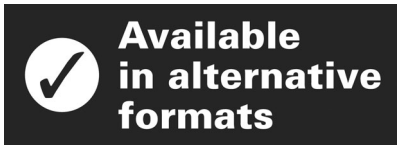
Research into the level of electoral registration and turnout among South Asian communities in England and Wales.

Turnout of 59.4 per cent at the 2001 General Election was the lowest since 1918 and it has been widely assumed that electors from minority ethnic groups are less likely to vote in general elections than white electors. This research provides a reliable, nationally representative estimate of South Asian electoral participation, using:

- electoral registers marked up at polling stations on election day
- election results from the 2001 General Election
- the 2001 Census of population.

The research, which also looks at factors affecting participation, finds that turnout among South Asian communities in England and Wales is as high, and in some cases higher, than that of the rest of the population, especially in areas where participation is generally low.

JOSEPH ROWNTREE FOUNDATION



This publication can be provided in alternative formats, such as large print, Braille, audiotape and on disk. Please contact: Communications Department, Joseph Rowntree Foundation, The Homestead, 40 Water End, York YO30 6WP.
Tel: 01904 615905. Email: info@jrf.org.uk

Electoral participation of South Asian communities in England and Wales

Edward Fieldhouse and David Cutts

JOSEPH ROWNTREE FOUNDATION

The **Joseph Rowntree Foundation** has supported this project as part of its programme of research and innovative development projects, which it hopes will be of value to policy makers, practitioners and service users. The facts presented and views expressed in this report are, however, those of the authors and not necessarily those of the Foundation.

Joseph Rowntree Foundation
The Homestead
40 Water End
York YO30 6WP
Website: www.jrf.org.uk

First published 2007 by the Joseph Rowntree Foundation

ISBN 978 1 85935 545 9

A CIP catalogue record for this report is available from the British Library.

Prepared by:
York Publishing Services Ltd
64 Hallfield Road
Layerthorpe
York YO31 7ZQ
Tel: 01904 430033; Fax: 01904 430868; Website: www.yps-publishing.co.uk

Further copies of this report, or any other JRF publication, can be obtained from the JRF website (www.jrf.org.uk/bookshop/).

Contents

Executive summary

The reported research is based on an analysis of a large sample of marked electoral registers and the 2001 Census. As such, unlike most previous research into South Asian turnout and registration, it is free from response bias. The key findings from the analysis include the following.

- South Asian adults are less likely to be registered to vote than the rest of the population, though this is partly attributable to a larger proportion of the population being born outside of eligible countries.

- Muslim communities (including both South Asian and other Muslim groups) have lower rates of registration than South Asian non-Muslim communities before adjusting for ineligibility due to country of birth. After allowing for the ineligibility to register to vote due to being born outside the UK, the registration rate for both South Asian groups is approximately 93 per cent.

- Muslim adults (including both South Asian and other Muslim groups) are more likely to be registered in areas with larger Muslim populations, but the same pattern is not evident for other South Asian adults.

- Other factors affecting registration include the stability of the population within an area (i.e. the proportion of people living at the same address as one year ago), the level of homeownership and unemployment, and the social class profile.

- Registered South Asian electors are more likely to turn out to vote than non-South Asians. Registered Hindu electors are the most likely to vote of all the identifiable religious groups common in the South Asian electorate.

- Registered South Asian women, especially those who are Muslim, are more likely to vote than South Asian men.

- All the identifiable South Asian groups turn out in greater proportions in areas where they are most concentrated. This is particularly evident for Muslim electors. This might be a result of enhanced mobilisation effects in more diverse areas.

- Statistical models of turnout confirm that higher levels of turnout are not explained simply by the social composition of the different religious groups.

- Models also show that the most important factors affecting turnout among South Asian communities include homeownership, the size of the religious minority population in the local area and the marginality of the constituency. However, the factors affecting turnout for the population who were not South Asian were slightly different, with the only common factors in the statistical models being homeownership (which was positively associated with the turnout of all groups) and the degree of marginality of the constituency.

1 The problem of measuring turnout and registration in South Asian communities

Outline of report

This report explores whether actual turnout and registration among South Asian communities is higher or lower than the rest of the population at the 2001 General Election in England and Wales. Apart from providing separate turnout and registration estimates for different South Asian communities and the rest of the population, the report notes the importance of local and contextual factors on levels of registration and turnout among different South Asian communities.

Introduction

This research uses the 2001 Census in conjunction with information from marked electoral registers from the 2001 General Election to provide a unique analysis of electoral turnout and registration among Britain's South Asian communities. The turnout of 59.4 per cent at the 2001 General Election was the lowest since 1918. Improving turnout at elections and improving levels of registration are two of the foremost problems facing the Government and society. The Electoral Commission was set up precisely 'to increase public confidence in the democratic process within the United Kingdom – and encourage people to take part'.[1] It has been widely assumed that minority ethnic electors are less likely to vote in general elections than white electors. Furthermore, electoral participation is regarded as both an indicator of the integration of minority communities and the quality of the democratic system. However, existing research that attempts to provide ethnic or religion-specific estimates relies heavily on survey data, aggregate data or small-scale case studies. A report published by the Electoral Commission shows that such data are highly unreliable when measuring turnout, particularly among minority groups.

We adopt an innovative approach to estimating turnout, which employs marked electoral rolls, election results from the 2001 General Election and the 2001 Census of Population. This allows us to make the most comprehensive and reliable, nationally representative estimates of South Asian electoral participation in Britain, and the social and political factors affecting it. This study represents the first large-scale, nationally representative, systematic analysis of actual (rather than reported) registration and turnout among South Asian communities.

The timing of the research is critical. Voter engagement is high on the political agenda, and the availability of the 2001 Census of Population collected within little more than a month of the 2001 General Election provided a unique opportunity to make accurate estimates of both turnout and registration. The empirical results should inform debates about citizenship, the decline in participation in the formal democratic process, and alternative explanations of differences in turnout and registration across and within different South Asian communities.

Variations in turnout

Voter turnout in Britain is unevenly distributed, and varies between different social and demographic groups, and between geographical areas (Swaddle and Heath, 1989; Johnston and Pattie, 1998). In particular, minority ethnic groups are often identified as having lower levels of participation in the formal democratic process than the white population (Anwar, 1990; Saggar, 1998). However, there are substantial differences in turnout and registration between different minority ethnic groups. For example, people of Indian heritage have been found to have comparable (and sometimes higher) rates of turnout than the white population.

At the area level, constituency turnout is related to a number of social and political factors including the class composition, the age profile and the electoral context. The proportion of electors from ethnic minorities has been found to be negatively associated with turnout. However, this ecological relationship does not necessarily hold at the individual level. Although ethnic minorities live in areas of lower than average turnout, their own levels of participation may be higher than an area-level analysis might suggest.[2]

Registration and turnout

> The turnout rates that are frequently reported are likely to be an under-recording of the number of people not voting in the UK. Reported figures do not take into account those people who are not registered to vote. (Purdam *et al.*, 2002)

Measures of turnout usually ignore the problem of non-registration. As many as 15 per cent of non-voters are not registered to vote and different sections of the population have differing levels of registration (Electoral Commission, 2001). For example, evidence suggests that, although parts of the South Asian communities

may have higher than average levels of turnout, levels of registration are generally lower than for the white population. According to data from the British Election Study, in 1997, 97 per cent of white and Indian voters were registered, whereas registration rates among Pakistani and Bangladeshi voters were 90 and 91 per cent respectively. A recent report by the Electoral Commission, based on a study conducted by the Office for National Statistics, estimated registration in 2000 to be approximately 93 per cent and found that, while black and minority ethnic communities as a whole had lower rates of registration than whites, the rates for South Asians were remarkably similar, with little difference between white Britons, Pakistanis and Bangladeshis (Electoral Commission, 2005).

Like turnout, registration varies geographically and is lowest in inner-city areas where ethnic minorities are most heavily concentrated. Because of the high levels of non-registration of eligible voters, particularly in certain areas and in certain groups, it is not possible to assess participation simply by reference to turnout. Registration must also be taken into account.

Measurement issues

Survey data on turnout within BME communities are generally inadequate. First, there is usually an insufficient sample to look at ethnic differences and, second, non-voting is widely under-reported. For example, a MORI survey taken shortly after the 2001 General Election showed turnout among white and Asian electors to exceed 80 per cent, compared to 70 per cent among black electors, when in reality turnout in the Election as a whole was only 59 per cent. The 1997 and 2001 British Election Survey (BES) used marked electoral registers to validate turnout among respondents, and shows large-scale discrepancies between reported turnout (and registration) and actual behaviour. There are various reasons for survey unreliability, including biased reporting of respondents and differential non-response to surveys. One alternative is area-based analysis of electoral returns but, as noted above, estimates for ethnic minorities are based on potentially spurious inferences from aggregate to individual data.

This research uses innovative methods to analyse levels of voter turnout and registration, focusing on differences within and between South Asian communities and the population more widely. We restrict our analysis of religion differences to the South Asian population, as software is available to distinguish the origin of Asian names on the electoral register.[3] Unlike previous research into minority ethnic participation, we will measure *actual* individual-level turnout using marked electoral registers without relying on reported turnout (as in sample surveys) or ecological

inference.[4] Non-registration is assessed using data from the 2001 Census. This research provides the only large-scale, nationally representative study of turnout *and* registration among British South Asian communities. The findings will inform and improve future estimations of turnout based on survey or ecological data.

Methods

This research uses marked registers from the 2001 General Election, for a sample of 97 wards, based on a stratified random sample (see Appendix 1 for more details). Using 1991 Census data, we stratified wards according to the percentage of the population which was South Asian. Wards were sampled disproportionately in areas with a large Asian population to ensure the effective coverage of different subgroups but weights are applied to make the sample nationally representative. All electors were included in the selected wards, which were used as the primary sampling units (see Appendix 1 for more details). Registration is assessed by comparing the Census population with our sample of marked electoral registers from the 2001 General Election. The marked registers are analysed using name recognition software (Nam Pehchan and SANGRA), which is able to identify names with a South Asian origin (i.e. from the Indian sub-continent).

In this report we will:

- provide accurate estimates of the level of electoral registration in the 2001 General Election
- provide accurate estimates of the level of turnout once non-registration has been allowed for
- provide separate such estimates of turnout for Muslim, Sikh, and Hindu communities and the rest of the population
- provide an improved understanding of local and contextual factors affecting levels of registration and turnout among different South Asian communities
- assess the reliability of ecological methods of estimating ethnicity.

2 Registration of South Asian populations: new evidence from the 2001 Census

Key findings

This chapter provides accurate estimates of registration among South Asian communities and the rest of the population at the 2001 General Election in England and Wales. Our analysis seeks to measure registration rather than explain it. However, we do show that ineligibility due to nationality plays a significant part, as does the geographical distribution of South Asian groups. For instance, we find that Muslim registration is highest where there are more Muslims and that a similar pattern, albeit smaller, exists for non-Muslim South Asians. In measuring these factors we also provide some insight into other factors (unemployment and homeownership) associated with registration.

Introduction

Although most policy debate about electoral participation concerns improving turnout, a substantial minority of the adult population never even reach it as far as the electoral register, let alone the ballot box. As a result, the statistics on which these debates are based may be misleading. The accuracy of reported levels of turnout is directly related to the completeness of the electoral register and estimates of participation based on the turnout of registered electors tend to overstate real turnout levels. In some countries, notably the United States, estimates of turnout are routinely based on the voting-age population or VAP, although, since 2001, the voting eligible population or VEP estimate is now used by a number of leading US scholars (see McDonald and Popkin, 2001). In most European countries, the denominator for turnout calculations is the registered electorate, which can be as much as 7 per cent higher than the VAP (e.g. 2000 Spanish parliamentary elections).

Obtaining reliable registration rates can be a difficult and imprecise process given uncertainty about the size of the eligible voting-age population (because of census undercoverage, temporary residency of foreign nationals, etc.). In particular, research that attempts to provide ethnic- or religion-specific estimates relies heavily on survey data, aggregate data or small-scale case studies. Most surveys focus on turnout rather than registration and in any case struggle to overcome the problems of

misreporting, non-response bias and a small sample size. Some surveys partly overcome these problems by validating whether people really voted using the marked electoral register and also by providing booster samples of minority ethnic electors (e.g. the 1997 BES 'black and minority ethnic' booster sample). However, despite the undoubted value of such surveys, they often suffer from a small sample size. For instance, the 1997 BES 'black and minority ethnic' booster sample contains only 227 Asians of Indian origin and 124 Asians of Pakistani origin. An alternative approach is to use area-level (or geographical) relationships between the size of the minority ethnic population and the level of turnout. However, as noted above, this approach is based on potentially spurious inferences from aggregate to individual data. In short, recent research seems inconclusive in assessing registration, particularly for different South Asian groups.

In this chapter we use information from the complete sets of marked electoral registers for a sample of 97 electoral wards at the 2001 General Election in conjunction with the 2001 Census of Population in order to estimate levels of registration in South Asian communities. These rates are used in the following chapter to calculate revised estimates of turnout based on the VAP. In 2001, the General Election (7 June) and Census Day (29 April) were remarkably close. The close co-incidence of an election and a census provides a unique opportunity to undertake analysis of registration as well as turnout at a time when voter apathy was a key election issue.

Registration is assessed by comparing the census population with our sample of marked electoral registers from the 2001 General Election. These are analysed using name recognition software, which is able to identify names with a South Asian origin (i.e. from the Indian subcontinent). Together with geographical population information from the 2001 UK Census, this information allows a unique analysis of electoral turnout and registration among Britain's South Asian communities.

Registration: an overview

For reasons of scrutiny and legitimacy, it is a key requisite of western democracies that a citizen must be registered to vote before he/she can participate in elections. In some countries (Belgium, Denmark, Netherlands, Spain) this takes the form of a national citizens' register, while others (Australia, France, Germany and the United Kingdom) use a voters' or electoral register. Yet whether a citizen can be on a list of registered voters varies between and within countries. For instance, in Nordic countries (Denmark, Sweden and Norway), any foreign person who has been resident in the country for more than three years has the right to vote, while the

figure is five years in the Netherlands. Some countries also require citizens to register by appearing at a registration office (France), while others compile and update registers through a combination of mail, door-to-door registration (Germany) and even the internet (Australia). The majority of countries update voter registration either continuously or annually, although some countries (Italy and Japan) register voters periodically, often just before an election.

Box 1 UK registration: an overview

- UK voter qualification age is 18.
- Registers are compiled by local authorities, which write annually to residents and request the completion of a form.
- The electoral register includes all those in a household who are aged 18 or over, as well as those 17 year olds who will become eligible to vote during the lifetime of the register.
- Under UK electoral law, registration is open to British, Irish or Commonwealth citizens, or members of a European Union state.
- British citizens living abroad can register as an overseas elector and are eligible to vote in UK and European parliamentary elections for up to 15 years after they left the country.
- Rolling registration was introduced at the 2001 General Election. The system allows individuals to update their details during a particular year.
- The register is now updated each month, apart from during the annual canvass period (September, October and November), and people can register to vote in the weeks before the election, but not once the election has been called.
- For the 2001 General Election on 7 June, new electors were required to register before 5 April. This led to a 1.3 per cent increase in the number eligible to vote in 2001 compared to 1997 (Electoral Commission, 2001).

In the majority of European countries (excluding France and Ireland), registration is compulsory, although the law is implemented to varying degrees. This has been frowned on by some in the United States as an abuse of individual civil liberties. In the United States, every state apart from North Dakota has voluntary registration procedures where the emphasis is on the citizen to register. Despite recent efforts to facilitate an increase in registration levels by linking driver licence renewal with voter registration (1993 National Voter Registration Act or ‘Motor-voter’ law), the US

continues to have much lower registration rates than its European counterparts. For instance, the US Bureau Current Population Survey estimated registration of the voting-age population in the November 2002 congressional elections at 69 per cent. That said, around 10 per cent of French citizens persistently abstain from registering for national elections (Mény, 2002).

Previous research findings

In the UK, evidence from comparing the 1991 Census and the Post-enumeration Survey estimated that 7.1 per cent of the people eligible to vote were not on the electoral register (Smith, 1993). A later study estimated that 4.8 per cent of people enumerated in the 1991 Census were not on the electoral roll (Heady *et al.*, 1996), while, in 1992, out of 426 constituencies, nearly a quarter had an eligible electorate of 500 more than were found on the register and, in two constituencies, the difference was over 3,000 (Pattie *et al.*, 1996). At the 2001 General Election, one study estimated registration at just under 97 per cent (IDeA, 2002). Another recent attempt to determine UK registration rates used the final mid-year estimates for 2001 to estimate the populations of England and Scotland aged 18 or over by June 2001 and subsequently concluded that registration levels were 97.0 per cent in England and 99.4 per cent in Scotland (Dorling, 2007, forthcoming). When compared to 2001 Census figures, it was estimated that around 7 per cent of people in England and Wales were not on the electoral register in 2002. However, the author admits that the figures should be treated with caution given the uncertainty of population estimates in London and the North West (Dorling, 2007, forthcoming). An estimate for 2004 suggests that UK registration rates might range from 92 to 93 per cent according to a study conducted for the Electoral Commission by the Office for National Statistics (ONS) (Electoral Commission, 2005). One of the strengths of the ONS survey is that it checked census and labour force survey records against the electoral register and is therefore likely to have a high degree of accuracy, although the sample sizes for minority groups were relatively small.

Electoral registration in Britain is unevenly distributed and varies between geographical areas (Smith, 1993) and between different social and demographic groups (Todd and Butcher, 1981; Smith, 1993). For instance, Smith (1993) estimated that non-registration rates were 2.2 per cent higher for men than women and found levels of non-registration to be higher for the youngest age groups (17 attainers and those in their early twenties) than for the 50 and over age group. In particular, substantial differences in registration rates have been identified between minority ethnic groups (Anwar, 1994, 1998; Smith, 1993; Saggar, 1998). For example, those of black African heritage often record the highest levels of non-citizenship, although

recent findings are less conclusive for different South Asian groups. Recent research, based on the 1997 British Election Survey (BES), which employed a 'black and minority ethnic' (BME) booster sample, found registration levels of 96.9 per cent for Indians, 90.2 per cent for Pakistanis, 91.3 per cent for Bangladeshis, 87.1 per cent for black Africans, 96 per cent for black Caribbeans and 96.9 per cent for whites (Saggar, 1998). Also, there were a substantial number of respondents who stated that they were registered to vote but not at the particular address at which the interview was conducted. Levels of British citizenship vary across BME groups with the highest levels of non-citizenship among black Africans and Bangladeshis. There was no similar booster sample in the 2001 BES. More recently, a face-to-face sample survey across five local authority areas found non-registration levels to be higher among Indians (24 per cent), black Caribbeans (26 per cent) and black Africans (25 per cent) than among whites (18 per cent), Pakistanis (17 per cent) and Bangladeshis (13 per cent) (Anwar, 1998). The Electoral Commission/ONS study discussed above suggested that the percentage not registered in South Asian communities was much lower: 6 per cent for Indians and Bangladeshis, and 8 per cent for Pakistanis, compared to 17 per cent for all BME groups and 6 per cent for whites. The study also found that there was a strong relationship between non-registration and nationality, which is also reflected in our results (see below).[1]

Box 2 Why do registration rates vary for minority groups?

1 Variations may be dependent on the methods used by electoral registration officers and diverse local authority policies on updating the register (Smith, 1993; LGA, 2000).

2 Registration offices have not sufficiently changed their practice to meet the needs of the BME electorate (Anwar, 1990, 1998).

3 Language difficulties and unease about dealing with officialdom.

4 Concerns with anonymity and fear of harassment.

5 Doubts about residence status affect BME communities disproportionately more than the wider population and have contributed to varying levels of non-registration (Anwar, 1990, 1996, 1998). Survey evidence from Bradford found that deliberate non-registration among Asians was much lower than other BME groups (Le Lohe, 1990).

While we do not set out to explain registration, but rather to measure it, we suggest that nationality does play a significant part, as does the geographical distribution of South Asian groups. In measuring these factors we also provide some insight into other factors associated with registration.

Constituency-level registration estimates

Calculating registration rates at the aggregate level

After obtaining the registered electorate for each constituency for England and Wales, we derived the voting-age population (VAP) from 2001 Census data (see Appendix 3). Initially, we examined registration rates for the country as a whole and by parliamentary constituency.

Using these data we estimate that in England and Wales there were 40,314,816 people who were eligible to participate in the General Election (VAP), whereas only 39,205,725 people were registered to vote. The estimated registration rate for England and Wales was 97.25 per cent. Not surprisingly, there were wide spatial variations in estimated registration rates.

Table 1 shows the top ten constituencies with registration rates above 100 per cent; in other words more adults were registered to vote than there were adults in the population to register! This discrepancy could be explained by census underenumeration (the denominator) or by inaccuracies in the register (the numerator), including the failure of electoral registration officers to adequately update the register, students who are registered at a home address or even adults who are still registered in these constituencies but either live or work elsewhere. Nine of the highest ten constituency registration rates were found in the North West. Dorling (2007, forthcoming) also noted a 'clustering of possible underenumeration in the North West' and hypothesised, following an assessment of registration rates in Bolton and the Wirral at the previous election, that the electorates in these seats, for one reason or another, were probably inflated.[2] However, Dorling (2007, forthcoming) stresses that population estimates in the North West were also probably a little low.

Table 1 The ten highest constituency registration rates in 2001

Constituency	VAP	Registered electorate	Registration estimate
Birkenhead	56,851	69,726	106.82
Dorset West	71,001	74,016	104.25
Wirral South	58,187	60,653	104.24
Liverpool Wavertree	69,605	72,555	104.24
Bolton North East	66,716	69,514	104.19
Manchester Blackley	56,989	59,111	103.72
Macclesfield	70,631	73,123	103.53
Wirral West	60,251	62,294	103.39
Liverpool Walton	64,186	66,237	103.20
Manchester Central	64,226	66,268	103.18

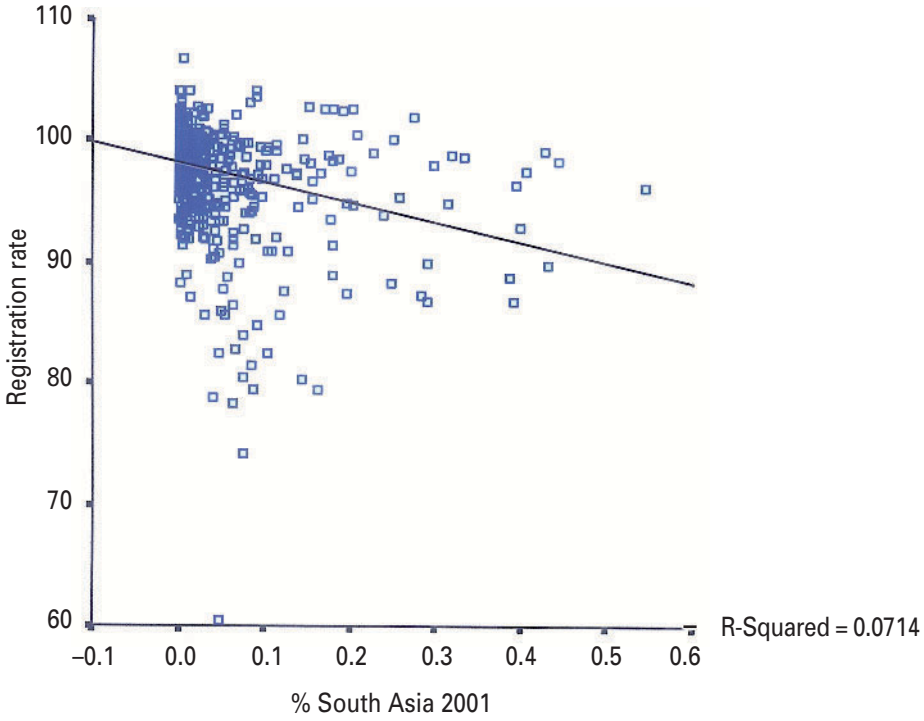
The lowest registration rates in England and Wales were found exclusively in London (see Table 2), including areas with large South Asian populations (e.g. Brent East). In total, 32 constituencies recorded registration rates below 90 per cent, while six have more than 20 per cent of the VAP not included on the register. It is possible that these constituencies, particularly in Central London, include a number of adults who are simply not registered to vote (registered at their home address) and that such areas could contain fewer households than the Census states (see Dorling, 2007, forthcoming).[3] These constituencies are likely to have an extremely mobile population, making estimates difficult to ascertain.

Table 2 The ten lowest constituency registration rates in 2001

Constituency	VAP	Registered electorate	Registration estimate
Kensington and Chelsea	102,392	62,007	60.56
Cities of London and Westminster	96,991	71,935	74.17
Hampstead and Highgate	83,268	65,309	78.43
Hammersmith and Fulham	100,515	79,302	78.90
Brent East	73,031	58,095	79.55
Ealing, Acton and Shepherds Bush	88,814	70,697	79.60
Holborn and St Pancras	78,104	62,813	80.42
Tottenham	81,433	65,567	80.52
Regent's Park and Kensington	93,106	75,886	81.50
North Southwark and Bermondsey	89,074	73,527	82.55

In order to explore whether there was any connection between areas of low registration and areas with large South Asian populations, we estimated the correlation at constituency level. Figure 1 illustrates the significant negative relationship (–0.267) between South Asians and registration. But does this ecological relationship hold at the individual level? While South Asian electors may live in areas where registration is generally much lower than elsewhere, their own registration rates might be much higher. Only by using individual data from our sample are we able to ascertain whether such an ecological fallacy exists. This is explored below.

Figure 1 Nature of aggregate relationship at constituency level; plotting 2001 General Election registration rate against per cent South Asian (2001 Census data); correlation coefficient –0.267



Generally, the constituency registration rates provide a 'ballpark' guide to the level of registration across England and Wales. However, because they are based on aggregated official electorates, no adjustments can be made for the existence of ineligible electors on the registers. Not only do some of these constituency estimates suffer from probable inflation of the registered electorate or census population, but it is also impossible to gauge accurate registration rates among different South Asian communities at this level of geography. We therefore turn to estimating levels of electoral registration in 2001 for census output areas in our sample of wards. Output areas (OAs) are the smallest geographical areas for which 2001 UK Census data are released. They nest into wards and are built up from unit postcodes (Martin, 2002). The 2001 Census will provide population information for OA and ward by religion and ethnicity. The number of registered electors of South Asian and other origins are then compared with the relevant census population.

Sample ward registration estimates

As noted earlier, 33 of our 97 sampled wards were left unchanged following local government boundary changes. We obtain sample ward registration rates using a similar method as employed at the constituency level.[4] The numerator is the registered electorate from our sample wards. However, it has been adjusted to take account of data-ageing problems (deleted names that were originally part of the registered electorate) and those attainers who didn't reach the voting age by the date of the election. This was possible as we have the unaggregated electoral registers for our sampled wards. We did not deduct EU citizens who are only permitted to vote in local and EU elections since these legitimately appear on the register even if they were not entitled to vote in the General Election.

The registered electorate is divided by the VAP to obtain ward registration rates. The overall registration rate across the 30 wards is 94.4 per cent. As at the national level, there are spatial variations in registration with a number of wards (St Nicholas and Longford) recording non-registration rates of less than 1 per cent, while others (Burngreave and Bradford Moor) have in excess of 10 per cent of citizens not registered. There is also evidence of significant within-constituency variation in registration rates. While two of the three Oldham East and Saddleworth wards recorded similar registration levels, the non-registration rate in the other ward was 6 per cent higher. There were also wide disparities between those wards that recorded the lowest registration rates and the estimated constituency registration rate of which the ward is a part (see Table 3). In some cases the variation was as much as 20 per cent.

Estimating South Asian registration: using output areas

Box 3 Terminology

Numerator

Number of registered electors from our sample of marked registers.

Denominator

Voting-age population (VAP) derived from census output areas. The VAP is amended to take account of deaths, attainers and ineligibility due to birthplace.

Table 3 Sample ward registration rates

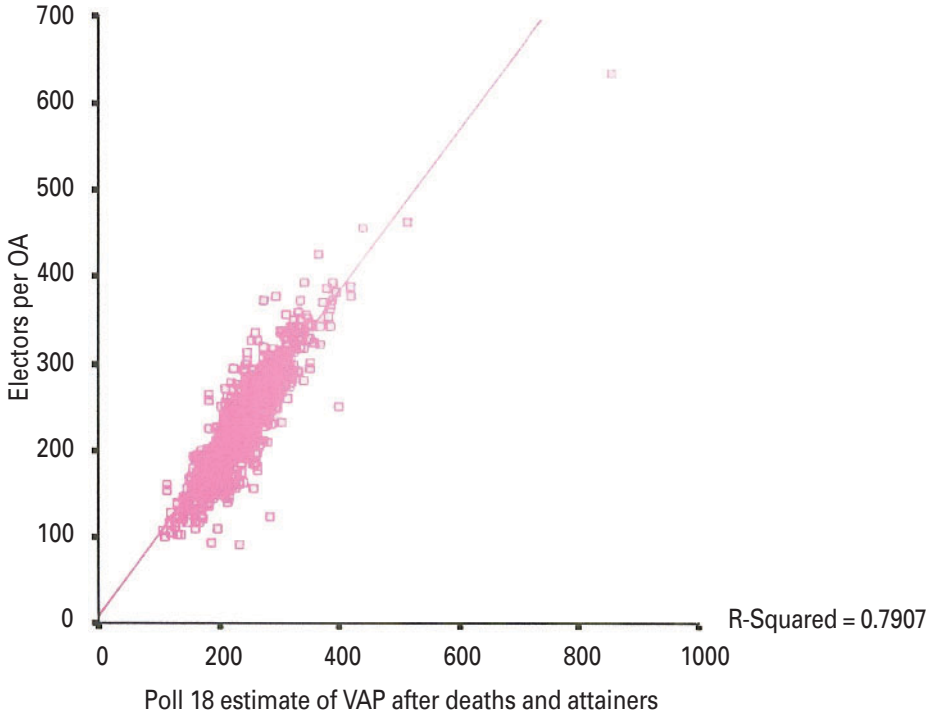
Sample wards	VAP (nos)	Registered electorate (nos)	Registered electorate/ VAP (%)
St Gabriel's	4,550	4,053	89.1
Darton	10,909	10,771	98.7
St James	6,956	6,332	91.0
Saddleworth West	9,062	8,964	98.9
Royton North	8,467	8,394	99.1
Crompton	8,840	8,561	96.8
Woodhouse Park	8,118	7,884	97.1
Hertford Heath	2,272	1,759	77.4
Great Sankey South	7,851	7,726	98.4
Weybridge South	3,194	3,039	95.1
Wheatley	8,412	6,676	79.4
Martin's Wood	4,724	4,723	100
Shaw	8,288	8,035	96.9
Wakefield North	11,866	11,232	94.7
Brockmoor & Pensnett	10,720	10,412	97.1
St Nicholas	4,506	4,555	101.1
Longford	7,379	7,414	100.5
Selly Oak	22,257	20,700	93.0
Burngreave	10,389	8,578	82.6
Pillgwenlly	3,903	3,837	98.3
Upper Stoke	13,106	12,616	96.3
St Thomas's	9,536	8,326	87.3
Riverside (Cardiff)	9,705	9,443	97.3
Crosland Moor	12,180	12,002	98.5
Redwell East[a]	3,522	3,345	95.0
St Matthew's	9,038	8,714	96.4
Aston	18,279	17,084	93.5
Coldhurst	7,891	7,908	100.2
Sandwell	20,712	19,316	93.3
Bradford Moor	11,296	9,852	87.2

a Huntingdon West and Town are removed due to missing cases following photocopying error. Tyisha (Llanelli ward) VAP was 3,199, while the registered electorate was 1,584 – this ward may have been subject to redistricting. Redwell East only – final register includes the whole of Redwell (amended register = 5,983).

Estimates of the number of registered electors and information on whether they voted are derived from our sample of marked electoral registers. These were allocated to geographical areas using the All Fields Postcode Directory (AFPD) (see below). The census output area (OA) is the smallest level of analysis for which we can obtain estimates of population and registered electors. This is therefore the most appropriate level of analyses to estimate registration and examine geographical variations. As a unit of analysis it has the additional advantage that we can correlate registration levels with the population characteristics (taken from the Census) at a fine level of geographical detail. More importantly, OAs also provide the building blocks to generate an aggregate VAP/VEP to compare with our sample of registered electors (see Appendix 3 for a fuller discussion of the methods).

Figure 2 shows the relationship between the number of electors and the size of the VAP (the denominator) for valid OAs in our sample. Valid OAs are defined as all those where the number of residential postcodes identified in our sample and matched to an OA exactly matches the number of residential postcodes in the AFPD.[5] It is clear from Figure 2 that there is a very close relationship between the two numbers, as would be expected. However, there are departures where some OAs have substantial differences between the two estimates. This may be due to low registration in some areas or may be due to inflated registers in others. By aggregating or summing across all the areas for which we have valid data, we are able to achieve an accurate estimate for England and Wales (see below). This analysis provides a superior method to most aggregate approaches, as it includes adjustments to both the denominator and numerator, as well as allowing us to disaggregate by religion. Full details of the method used along with how we identified South Asian electors can be found in Appendices 2 and 3.

Figure 2 Size of registered electorate against size of voting-age population



2001 South Asian registration

Box 4 Religious groups used in the calculation of registration

1 Non-South Asian

2 Muslim (all Muslims, plus other Pakistanis and Bangladeshis with no recorded religion)

3 South Asian non-Muslim (all Asian or mixed white and Asian Hindus and Sikhs, plus Indians, Pakistanis and Bangladeshis who are not Muslim)

4 All South Asians

Table 4 shows the estimated registration rates for all our sampled OAs (unweighted) by the identifiable religious groups, before and after adjustments for country of birth (COB). It also includes weighted registration estimates for England and Wales. Only areas where the denominator for a group is greater than 6.5 are included in the estimates of that group, as small census cells were subject to rounding for statistical disclosure control reasons, making them unreliable (Rees *et al.*, 2005).[6]

Table 4 Registration by religion/ethnicity, (a) unweighted (sample only) and (b) weighted for sample design – based on all output areas where denominator is greater than 6.5 (England and Wales only)

Registration by religion/ethnicity	With COB adjustment (unweighted, %)	Before COB adjustment (unweighted, %)	With COB adjustment (weighted, %)	Before COB adjustment (weighted, %)
Overall (1,823)	100.6	96.8	100.4	98.5
Non and other Asian (1,823)	101.5	98.1	100.5	98.9
Non-Muslim South Asian (763)	92.6	90.6	90.7	88.1
Muslim (944)	96.4	89.5	91.5	82.0
All South Asian (1,182)	95.5	90.7	93.4	86.2

Note: the number of valid OAs is shown in brackets.

The adjusted figures assume that no persons born outside the UK, Europe or the Commonwealth were eligible to vote. The figures in the third and fifth column do not make this adjustment and assume all persons of voting age are eligible. Naturally the unadjusted rates are lower, since the adjustment involves removing people born outside of eligible countries from the denominator. While the unadjusted figures may understate registration somewhat (due to the existence of genuine ineligibles), they may provide as reliable an estimate of registration as the adjusted rates, since many persons born in ineligible countries are naturalised or enjoy dual citizenship.

Unfortunately, there is no basis on which to estimate the proportion of this population who are eligible to register to vote (Electoral Commission, 2005). However, for the Muslim population in particular, it is important to take the number of people born outside of eligible countries into account since they constitute a large proportion of the Muslim population. This inevitably affects registration rates. For example, the Electoral Commission research showed that non-registration among Muslims living in the UK for ten years or more was only 6 per cent compared to 14 per cent among all Muslims.

Table 4 shows that, before making any adjustment for country of birth, the lowest unweighted rates of registration in our sampled wards are for Muslims followed by non-Muslim South Asians. Both groups have considerably lower rates than the non-Asian population. However, once country of birth has been taken into account, weighted and unweighted rates are considerably higher and the differentials are smaller. Indeed, the unweighted Muslim rate for our sample is above 96 per cent, higher than the non-Muslim South Asian rate of approximately 93 per cent. The non-Asian unweighted rate after country of birth has been adjusted exceeds 100 per cent suggesting that the adjustment is removing too many people from the denominator. This is not surprising since some of those born outside of eligible countries will be naturalised and eligible to vote. In addition both sets of estimates may be partly inflated by redundancy in the register or by census undercount. This large discrepancy in the unweighted Muslim rate reflects the greater number of Muslims counted in the Census who are born outside of eligible countries (e.g. in North Africa and South East Asia). While we are confident that we have identified the vast majority of Muslims in the electorate, both South Asian and from other parts of the world, there is likely to be a large number of Muslims who are not eligible to vote and hence would not be expected to be on the register. Indeed, if we take the non-adjusted rates as the baseline estimate, a substantial proportion of the difference between South Asians and the rest of the population is accounted for by differences in country of birth.

For the mainly Hindu and Sikh ‘other South Asian’ groups, the impact of country of birth is smaller than for Muslims, as this group is predominantly either UK or Commonwealth born. The all South Asian registration estimates, both unweighted and weighted, are based on a larger sample of OAs than both religious subgroups (more OAs where the denominator is greater than 6.5), hence the higher overall South Asian registration rates in two of the four columns in Table 4.

The effect of the proportion of people within each religious group who are born outside of the specified eligible countries on the registration rate for that group can be illustrated by a simple regression analysis. The dependent variables are the OA

registration rates for each religion category and the explanatory variables are the per cent within each religious group who were born in ineligible countries (i.e. the 'religion-specific' rate of ineligibility). The results show there is a decrease in registration across all groups as the percentage born outside eligible countries increases, and that the rate of decrease is smaller for those who are not South Asians. In other words there is a significant negative relationship between South Asian (Muslim, non-Muslim South Asians and all South Asian) registration and ineligibility due to birthplace. While, in the absence of a reliable estimate of the proportion of people born outside of eligible countries who are naturalised and eligible to vote, it is reasonable to report the unadjusted rate, it is important to bear in mind that non-registrants clearly include many who are ineligible.

The discussion above relates to areas included in our sample. However, because we used a stratified sample, making inferences about England and Wales as a whole is not straightforward. Simple stratification weights proportional to the sampling fraction for each stratum can be applied, though these introduce a potential secondary problem. In areas with very small South Asian populations, any errors in either the numerator (e.g. misclassification) or the denominator (e.g. census underenumeration) will have a disproportionately large effect on registration rates. These areas also have the highest weights as they have the lowest sampling fractions (see Table A1.1), meaning weighting will exaggerate any such errors. Though this is not a problem if errors are distributed equally in both directions, any systematic bias in errors could bias the overall weighted rate. As it happens, the estimated South Asian registration rates in these areas are lower than the rates for other areas (see Table 5 below), and therefore the use of weights has the effect of reducing the overall estimates of registration for South Asians. We cannot completely rule out the possibility that this effect is spurious (i.e. that rates in areas with small denominators are underestimated).

The resulting weighted figures for England and Wales are reported in Table 4 above. As explained, the rates are all lower than the unweighted rates, especially for Muslims (91.5 per cent after allowing for country of birth compared to 96.4 unweighted). The equivalent rate for other South Asians is just under 91 per cent. The overall South Asian rate is slightly higher than either subgroup separately due to inclusion of a larger set of valid OAs (the Muslim and non-Muslim rates for the 1,150 OAs used in the calculation of all South Asian were 92.8 and 93.0 per cent respectively). As explained above, the differences between these and the unweighted rates are due to the large stratification weight associated with the mainly non-Asian areas that have lower levels of registration for these groups. The relationship between the geographical concentration of Asian populations and the rate of registration is explored in more detail in the next section.

Geographical variations

Above we showed a negative constituency-level correlation between levels of registration and the size of the South Asian population. However, it was possible that this could have been the result of an ecological fallacy rather than lower registration rates of South Asians. In other words it might have been due to lower registration rates among the non-South Asian population. The disaggregated analyses in the preceding section dispelled this possibility. However, this does not mean that there were not geographical effects occurring whereby areas with larger South Asian populations experienced lower registration among South Asians and other voters alike. For example, this might be due to the concentration of South Asians in poorer neighbourhoods. As noted above, there may also be disproportionate measurement error in stratum 1.

Table 5 breaks down the registration rate, comparing both with and without country of birth (COB), of each group by the stratum in which they were sampled. As noted above, stratum 1 has the smallest proportion of South Asians (less than half a per cent) and stratum 5 the largest (more than 20 per cent). The table shows a very strong relationship between the size of the South Asian population (as represented by the stratum) and the levels of registration for South Asian groups. For both Muslims and non-Muslims, South Asian registration increases progressively with the size of the South Asian population, except for non-Muslims in stratum 5. As expected, when we make adjustments for those born outside of eligible countries, the rates are higher across the board, especially for Muslims. Indeed, the adjusted Muslim rate in stratum 5, where South Asians make up more than 20 per cent of the population, is nearly 98 per cent.

Box 5 Key points

- The larger the South Asian community, the better mobilised and the more politically engaged they become.
- Registration appears to be affected by belonging to a ‘religious enclave’ and therefore the potential mobilising affect of living in cohesive communities.
- The relatively isolated are more likely to be excluded from the democratic process. This is particularly true for Muslims, who, in areas with the largest South Asian populations, are more likely to be registered than non-Muslim South Asians.

Continued overleaf

- The data do reveal methodological concerns. South Asian rates for strata 1 and 2 are sufficiently low to arouse suspicion. These estimates are likely to be unreliable, as they are based on only a relatively small number of OAs.
- However, there is seemingly a direct and consistent relationship between South Asian population concentration and registration.
- It seems entirely plausible that rates in those areas sparsely populated by South Asians do indeed have low registration rates for those communities.

Table 5 Stratum percentage registration rates without and with country of birth adjustment based on all output areas where denominator is greater than 6.5

Registration	Stratum 1	Stratum 2	Stratum 3	Stratum 4	Stratum 5
Overall					
Without COB	99.1 (283)	99.3 (237)	96.8 (359)	95.3 (462)	95.9 (482)
With COB	100.1 (283)	101.4 (237)	101.3 (359)	100.2 (462)	100.4 (482)
Non and other Asian					
Without COB	99.1 (283)	99.4 (237)	97.4 (359)	96.8 (462)	98.4 (482)
With COB	100.1 (283)	101.3 (237)	101.3 (359)	101.1 (462)	103.8 (482)
Non-Muslim South Asian					
Without COB	58.6 (2)	75.9 (24)	83.3 (111)	91.6 (260)	90.9 (366)
With COB	86.5 (1)	79.0 (23)	86.1 (108)	93.5 (259)	92.9 (363)
Muslim					
Without COB	45.7 (13)	67.5 (25)	77.2 (154)	85.5 (341)	93.1 (411)
With COB	50.4 (11)	81.7 (16)	94.5 (128)	95.3 (334)	97.5 (407)
All South Asian					
Without COB	55.1 (18)	76.0 (59)	84.1 (230)	88.1 (411)	92.5 (464)
With COB	63.2 (15)	89.4 (49)	95.4 (214)	95.5 (409)	95.7 (463)

Note: valid OAs are shown in parentheses.

The relationship between religious composition and turnout can be seen in much more detail at the OA level. Figure 3 shows a clear relationship between South Asian Muslim registration and the proportion of the OA population that group makes up. Although there is a lot of variance where the Muslim electorate is very small, this is simply because many of those observations are based on very small numbers. The upward trend moving along the x-axis strongly suggests that registration is affected by belonging to a 'religious enclave' in the Muslim population. This could possibly be accounted for by enhanced community networks or social capital, and mobilisation, since it is in areas where Muslims are most densely populated that these effects would be expected to be most powerful. The picture for non-Muslim South Asians is very similar. Figure 4 shows a fairly clear and strong relationship between the percentage of the population made up by non-Muslim South Asians and their registration rate.

Figure 3 Muslim electorate and registration (by output area)

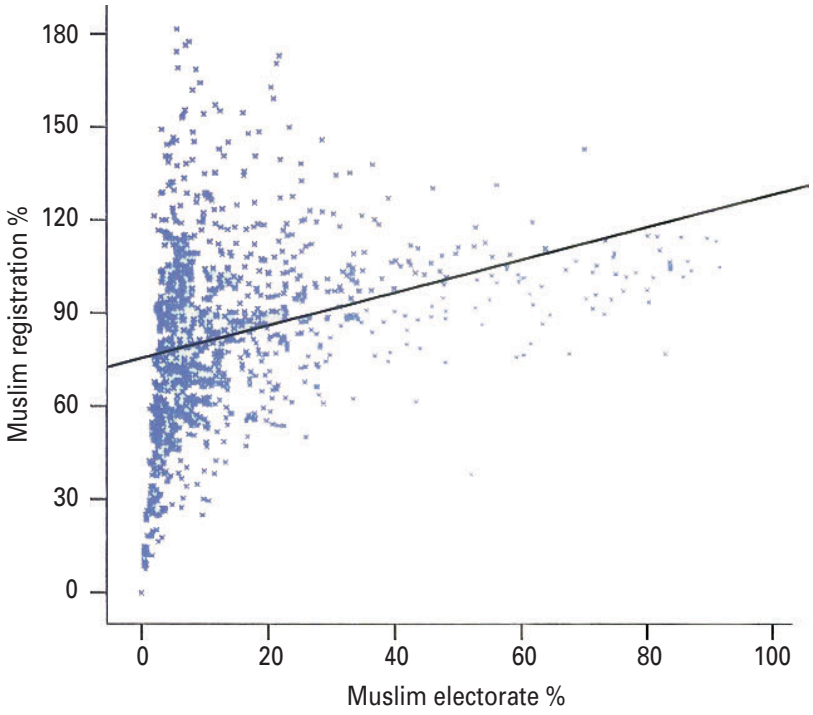
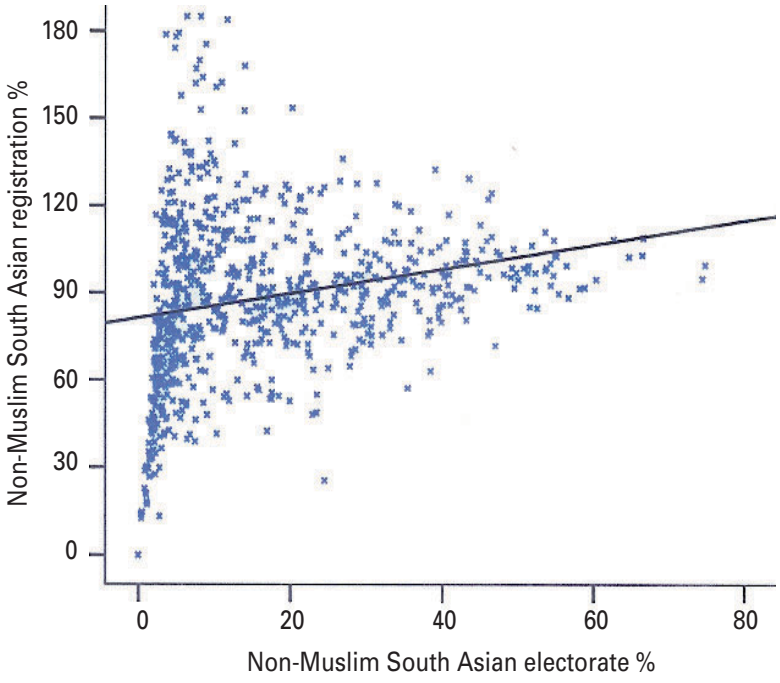


Figure 4 Non-Muslim South Asian electorate and registration (by output area)



We can confirm this relationship using a series of simple bivariate ordinary least squares (OLS) regression models (Table 6).[7] This confirms the positive relationship between Muslim registration and the number of Muslims living in an area. A similar statistically significant pattern exists for other South Asians and for all South Asians. The results suggest a clear positive association between where South Asians live and registration in general. At the same time non-South Asian registration is not affected by the proportion of non-South Asians in the area. Whether or not these relationships arise from mobilisation or social capital effects we cannot prove here, but it is clear that registration of South Asians is higher in the areas where those communities are most concentrated.

Table 6 OLS regression model of OA registration and composition of the OA population (weighted for number of size of denominator in OA)

Variable	All persons b coefficients	Non-South Asians b coefficients	Muslim b coefficients	Non-Muslim South Asians b coefficients	All South Asian b coefficients
Constant	97.40	98.07	78.03	85.18	81.31
OA % Muslim	–	–	+0.30*	–	–
OA % Hindu and Sikh	–	–	–	+0.17*	–
OA % South Asian	–0.03*	+0.004	–	–	0.21*
R-squared	0.01	0.00	0.09	0.02	0.05

* *Significant at the 95 per cent level.*
– Not included.

To return to the quandary posed above concerning the impact of weighting on the overall estimates, the statistical significance of the relationship between South Asian population share and South Asian registration seems to lend support to the argument for taking at face value the lower rates in strata 1 and 2, and hence trusting in the weighted national rates (rather than unweighted sample rates) reported in Table 4 above. In other words the national rate of registration for South Asians is approximately 86 per cent but, once country of birth has been taken into account, this rises to 93 per cent.

So far we have demonstrated, not only that registration rates are generally lower for South Asian communities, but also that these are affected by ineligibility of large proportions of the population, and that rates are highly variable according to the religious composition of the area. To substantiate the latter finding we now test whether this might be explained by the percentage of ineligible voters in each group.

Table 7 demonstrates that the positive association between Muslim registration and the geographical concentration of Muslim communities survives when controlling for ineligibility due to birthplace. There is also a positive relationship between non-

Muslim South Asian registration and the number of non-Muslim South Asians living in an area, with the coefficient larger for non-Muslim South Asians than for Muslims. For both Muslims and non-Muslim South Asians, the percentage born outside eligible countries for that religious group also has an independent significant negative effect on registration. For South Asians as a whole, registration is affected by the proportion of South Asians in the area but this relationship is weaker than for the disaggregated analyses. As for the religion subgroups, there is a significant negative effect related to the percentage born outside eligible countries. There are smaller effects for non-South Asians and the overall population.

Table 7 OLS regression models of OA registration (unadjusted for COB), controlling for country of birth

Variable	All persons b coefficients	Non-South Asians b coefficients	Muslim b coefficients	Non-Muslim South Asians b coefficients	All South Asian b coefficients
Constant	100.23	98.94	88.87	90.07	95.53
OA born in ineligible countries (% religion specific)	–0.97*	–0.39*	–0.78*	–2.68*	–1.55*
OA % Muslim	–	–	+0.16*	–	–
OA % Hindu and Sikh	–	–	–	+0.21*	–
OA % South Asian	0.01	+0.04*	–	–	+0.06*
R-squared	0.14	0.02	0.13	0.08	0.15

* *Significant at the 95 per cent level.*
– Not included.

To substantiate these findings we now test whether this might be explained by the socio-economic composition of the areas, rather than the religious composition. Again, we use simple linear regression models of registration while controlling for ineligibility due to birthplace and a number of socio-economic indicators (see Table 8).

Factors affecting registration

The variables used in the model (Table 8) include social and demographic variables that measure characteristics for that religious group (Muslim or non-South Asian Muslim, where available) and also characteristics of the area as a whole. The details of variables used in the model are included in Appendix 5. Table 8 shows the results of the analyses.[8] It is notable that most of the variation in registration is not accounted for by the independent variables in the model (reflected in the R-squared). However, a number of interesting findings do emerge.

Table 8 OLS regression models of OA registration (unadjusted for COB), controlling for socio-economic indicators

Variable	All persons b coefficients	Non-South Asians b coefficients	Muslim b coefficients	Non-Muslim South Asians b coefficients	All South Asian b coefficients
Constant	80.21	63.91	91.74	86.88	99.71
OA born in ineligible countries (% religion specific)	–1.17*	–0.41*	–0.85*	–2.48*	–1.31*
OA % Muslim	–	–	+0.14*	–	–
OA % Hindu and Sikh	–	–	–	+0.01	–
OA % South Asian	0.01	+0.03*	–	–	+0.01
Socio-economic variables					
OA unemployment (% religion specific)	–0.12*	–0.16*	–	–0.53*	–
OA owner occupation (% religion specific)	+0.04*	+0.04*	+0.16*	+0.07*	+0.04*
OA manufacturing (% all persons)	–0.10*	–0.07*	–	+0.45*	–
OA long-term ill (% all persons)	–	–	–0.63*	–	–0.56*
OA pensioners (% all persons)	+0.02*	+0.02*	+0.19*	–	+0.14*
OA high social class (% all persons)	–	–	–0.25*	–	–
Ward high social class (% religion specific)	–	–	–	–0.49*	–0.24*
Ward manufacturing (% all persons)	–	–	–	–	+0.30*
Ward agriculture (% all persons)	–	–0.60*	–	–	–
Ward full-time students (% all persons)	+0.20*	+0.23*	–	–	–
Ward non-migrants (% all persons)	+0.24*	+0.37*	–	–	–
Ward long-term ill (% all persons)	–0.17*	–	–	–	–
Ward car ownership (% all persons)	–	–	–	+0.46*	–
R-squared	0.20	0.08	0.18	0.16	0.18
Number of valid OAs	1,795	1,795	936	753	1,169

* *Significant at the 95 per cent level.*

– Not included.

Insignificant control variables were dropped from the models, with priority given to religion-specific variables over general variables and OA over ward.

Table 8 introduces the socio-economic control variables. Looking first at the overall registration rate, the number of people born outside of eligible countries remains a powerful negative influence. This is what we would expect given that we already know that those ineligible due to birthplace account for a large proportion of the

unregistered. In fact the model shows that, for every 1 per cent increase in those born in ineligible countries, there is just over a 1 per cent decrease in registration. The model also confirms that overall registration is not affected by the proportion of South Asians in the area. A number of the socio-economic and demographic controls are significant. For example, as we might expect, more stable population (the proportion living at the same address as one year ago) is positively associated with registration. It is well known that, when people move, there is often a considerable time lag before re-registering at the new address, thus bringing down registration levels. Registration is also positively associated with owner-occupation, older people and the number of students,[9] and negatively correlated with unemployment, manufacturing and the number of long-term ill (the latter at the ward level).

Table 8 also shows separate models for different religious groups since it is by no means necessary that registration for different groups should be subject to the same influences. For example, we showed above that the religious profile of the area was more important for some groups than for others. In these models we include variables that measure characteristics only for that group where possible, and also characteristics of the area as a whole.

The pattern for non-South Asians is fairly similar to the overall model, with a negative effect for the percentage born in ineligible countries and a small positive effect for the percentage South Asian. Both homeownership and the number of older people in the local area (OA) are positively associated with registration, while unemployment has a negative association, as does manufacturing. At the ward level, agriculture is negatively signed indicating that more rural areas tend to have lower registration. The number of full-time students also enters the model, at the ward level, and is (perhaps surprisingly) positively associated with registration. Still, the most important variables appear to be the percentage born in ineligible countries (negatively signed) and the number of non-migrants (positively signed).

For South Asians some interesting patterns emerge. After controlling for socio-economic composition, while the proportion of the religious group born outside of eligible countries considerably dampens all South Asian and both religion subgroup registration rates, there is no relationship between the density of South Asian population and registration, except for Muslims. The positive coefficient clearly indicates that Muslim registration is higher in the most Muslim areas and this is not attributable to social composition of those areas. However, this does not extend to non-Muslim South Asian areas, for which there is no religious compositional effect after controlling for socio-economic factors.

Social factors are also important to varying degrees, with South Asian and both religion subgroup registration rates significantly positively affected by homeownership for all groups. And the number of older people in the local area has a significant positive effect on both Muslim and overall South Asian registration. By contrast, the number of long-term ill, level of unemployment and high social class status are generally negatively associated with South Asian registration. At the ward level, car ownership, which tends to be correlated with general affluence, is positively associated with non-Muslim South Asian registration.

Conclusion

This chapter analyses the 2001 Census and a sample of marked electoral registers to estimate registration rates with a considerable degree of accuracy and provide comparative estimates for South Asian religious minorities and the rest of the population. Our key findings are as follows.

- At the constituency level there is a negative association between the size of the South Asian population and the level of registration.

- OA-level data disaggregated by religion derived from the electoral registers also show that South Asian adults are less likely to be registered than their non-Asian counterparts, although this can be partly accounted for by differences in country of birth.

- After allowing for ineligibility due to country of birth, the national (weighted) registration rate for both Muslim and non-Muslim South Asians is approximately 93 per cent.

- Statistical models demonstrate that, for non-South Asians and all South Asians alike, ineligibility due to birthplace remains the most significant factor influencing registration levels at the 2001 General Election in England and Wales.

- In areas where South Asian populations are more concentrated, rates of registration for South Asian electors are much higher. This relationship is easily missed in aggregate analyses.

- We found a strong relationship between levels of registration and the size of the Muslim population. Our models confirm this finding, with Muslim registration higher in Muslim areas even after controlling for social and demographic variables and for ineligibility.

- This relationship was observed for non-Muslim South Asians and all South Asians, though this did appear to be accounted for by the socio-economic composition of areas.

- Other factors that affect South Asian registration include the number of long-term ill, the level of unemployment, the number of older people and the extent of homeownership in the local area.

3 Turnout of South Asian electors: evidence from the marked electoral registers

Key findings

Using religious origin to aid comparisons with other data sources, the results in this chapter show turnout in 2001 is slightly higher (although not significant) for South Asian electors than for the rest of the population, but this varies by religious groups. Also South Asian turnout is significantly higher in areas where there are more South Asians in the electorate, which is where overall turnout rates are much lower.

Variations in turnout

While levels of participation in modern democracies continue to decline, turnout is increasingly seen as a key aspect of the accountability of governments and of citizenship. Turnout at the 2001 General Election (59.4 per cent) was at its lowest since 1918. This marked a dramatic fall since 1997 (71.6 per cent) and follows a period during which there was an underlying downward trend since turnout peaked in 1950 (Denver and Hands, 1997; Heath and Taylor, 1999; Clarke *et al.*, 2004). It barely recovered in 2005.

Voter turnout in Britain is unevenly distributed, and varies between different social and demographic groups and between geographical areas (Swaddle and Heath, 1989). In particular, minority ethnic groups are often identified as having lower levels of participation in the formal democratic process (Anwar, 1990; Ali and Percival, 1993). However, there are substantial differences in turnout and registration between different minority ethnic groups. For example, people of Indian heritage have been found to have comparable (and sometimes higher) rates of turnout than the white population.

Box 6 Key research

1997 British Election Survey (BES): key findings

The 1997 British Election Survey (BES), which employed a 'black and minority ethnic' (BME) booster sample, showed the following turnout rates (Saggar, 1998):

- 82.4 per cent for Indians
- 75.6 per cent for Pakistanis
- 73.9 per cent for Bangladeshis
- 68.7 per cent for Black Caribbeans
- 64.4 per cent for Black Africans
- 78.7 per cent for white voters.

There was no similar booster sample in the 2001 BES.

2001 MORI survey: key findings

- Asian and white turnout rates were considerably higher than those of black electors.
- However, the survey massively overestimated turnout among all groups and must be treated with some caution (Purdam *et al.*, 2002).

2005 MORI/Electoral Commission Survey: key findings

- Turnout was higher among the main Asian groups (Bangladeshis, Pakistanis, Indians) than black electors. Mixed-race electors had the lowest turnout rate of all BME groups.

At the area level, previous research also shows that constituency turnout is related to a number of social and political factors, including the class composition, housing characteristics, age profile, and the electoral and tactical context (Denver and Hands, 1997; Johnston and Pattie, 1998). The ethnic profile was also found to be a significant factor, with larger minority populations negatively associated with turnout after controlling for other factors (Purdam *et al.*, 2002). However, it is noted that this ecological relationship does not necessarily hold at the individual level. Although ethnic minorities live in areas of lower than average turnout, their own levels of participation may be higher than an ecological model might suggest.

Indeed, although low voter turnout at an aggregate level may be associated with concentrations of BME communities, evidence at the level of the individual voter points towards higher levels of turnout among sections of the minority ethnic population, notably Indian Asians (see Box 6 above). This has been shown using survey data at a national level and in the context of one of the proposed case studies (Anwar, 1990; Le Lohe, 1990; Saggar, 1998). Furthermore, because turnout has a strong spatial dimension, we might expect South Asians to have lower levels of turnout, as they live in areas characterised by low turnout. For example, BME voters are relatively more likely to live in safe seats and in areas of economic deprivation (e.g. inner-city areas). The geographical distribution of the minority ethnic population and the characteristics of those areas may have an impact on levels of turnout. However, until now, we have not known the relative levels of turnout of BME and white voters within areas (i.e. whether low turnout is characteristic of a specific community or a specific area).

Measurement issues

Box 7 Measurement issues

Survey data on turnout within BME communities is inadequate because

1 There is usually an insufficient sample to look at ethnic differences.

2 Non-voting is widely underestimated for two reasons: biased reporting of respondents (i.e. people claiming to vote) and differential non-response to surveys (i.e. non-voters less likely to respond to surveys).

Example: MORI survey taken shortly after the 2001 General Election

- It showed turnout amongst white voters to be 80 per cent. Asian voter turnout exceeded 80 per cent, compared to 70 per cent amongst the black electorate.
- In reality, turnout in the 2001 General Election as a whole was only 59 per cent.

Example: 2001 British Election Survey (BES)

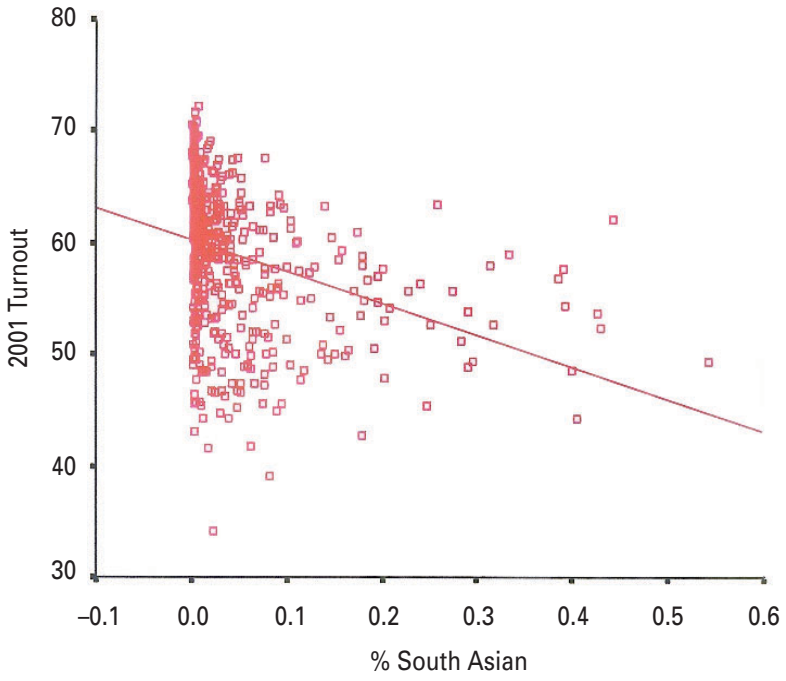
- It used the marked electoral registers to validate turnout among respondents and found large-scale discrepancies between reported turnout (and registration) and actual behaviour.

Continued

- BES 2001 turnout (weighted) was 71 per cent – 12 per cent below the actual turnout figure.
- Around 6 per cent was due to differential non-response bias while the other 6 per cent was due to misreporting.

One option is ecological (or area-based) analysis of electoral returns. However, as noted above, the main problem with ecological estimates of non-voting is that, while full population figures are reliable, estimates for ethnic minorities are based on potentially spurious inferences from aggregate to individual data (Robinson, 1950). Figure 5 illustrates the significant negative relationship (correlation = –0.333) between ethnicity and turnout at the area level. But does this ecological relationship hold at the individual level? While ethnic minorities may live in areas where turnout is generally much lower than elsewhere, their own participation rates might be much higher. Using individual data from our sample, we are able to address this ecological fallacy in more depth later in this chapter.

Figure 5 Nature of aggregate relationship at constituency level, plotting 2001 general election turnout against ethnicity (2001 Census data); correlation coefficient (–0.333)



Given that we are able to distinguish the origin of South Asian names by religion, as described in Chapter 2, marked registers are employed to ascertain the actual individual turnout of South Asians from our sample of wards in 2001 (see Table 9). Analyses are weighted to reflect sample design and the national turnout rate. Although in Table 9 we report results using the classification derived from Nam Pehchan and SANGRA separately, we focus on the combined classification reported in column 4, since this represents what we consider to be the most complete classification (see Appendix 4). A number of important findings emerge.

Table 9 Turnout rates (excludes postal voters) – weighted to be nationally representative

Religion/ethnicity	Nam Pehchan % voted corrected for national turnout	SANGRA % voted corrected for national turnout	Combined % voted corrected for national turnout
Hindu	61.7	61.1	61.3
Muslim	58.7	59.0	58.5
Sikh	60.7	58.7	59.7
Religion not determined	56.8	–	–
Other South Asian	–	57.4	55.8
All South Asian	58.9	59.5	59.4
Non-Asians	58.4	58.3	58.3
Total	58.4	58.3	58.3

2001 South Asian turnout

First, Table 9 shows South Asian turnout (59.4 per cent) was one percentage point higher than non-Asians (58.3 per cent).[1] Weights were applied to make the findings nationally representative. This possibly represents the most reliable estimate of South Asian electoral participation in Britain to date, and notably contrasts with survey estimates suggesting lower levels of turnout than their white counterparts. Second, South Asian turnout varied among religious groups. Hindus recorded the highest turnout in 2001, which is significantly higher than the overall rate (58.3 per cent). A higher percentage of Sikhs (59.7 per cent) also voted than non-Asians, while Muslim turnout was almost identical to non-Asian turnout. However, these differences are not significant at the 95 per cent confidence level.

These results appear to confirm recent survey findings that people of Indian heritage (predominantly Hindu and Sikh) have the highest level of turnout of all ethnic groups in Britain (Anwar, 1990; Le Lohe, 1990; Saggar, 1998). Yet previous survey evidence suggested that people of Muslim heritage were less likely to vote than non-Asians. Our findings suggest otherwise; by religion, turnout was around 0.3 percentage points higher than (and not significantly different from) non-Asians.

South Asian turnout by gender

Apart from language and religion, Nam Pehchan (but not SANGRA) identified South Asian names by gender. Validated estimates of turnout from the 2001 BES suggest that men and women voted in equal proportions. Yet there were apparently marked differences in reported turnout between minority ethnic men and women in 2001, with the latter far less likely to participate than their male counterparts (Norris *et al.*, 2004). Our evidence contradicts this.

Table 10 records the percentage who voted by religion and gender using the Nam Pehchan classification. Figures are provided for men, women and where gender was not determined by name. It is clear from looking at the sample sizes that the vast majority of those in the 'gender not determined' category were women. The results contrast with the national picture and also some previous survey-based estimates. Notably, turnout among South Asian women (64.6 per cent) was more than six percentage points higher than men (58.2 per cent). Muslim and Hindu women were the most likely to vote, and the rate for Muslim women exceeded that for men by over 7 per cent. Unfortunately, the data for Sikhs may be slightly misleading; given that many Sikh names are common for both men and women, it is not surprising that the vast majority of identified Sikh voters were placed in the 'gender not determined' category.

Table 10 Percentage voted by religion and gender (design and vote weight – Vgweight)

Religion	Gender not determined	Female	Male	Total	*n*
Hindu	60.0	65.7	62.5	61.3	152,099
Muslim	58.4	64.5	57.0	58.5	310,447
Sikh	59.7	57.9	60.1	59.7	91,712
Other South Asian	61.5	59.2	60.7	55.8	5,656
All South Asian	59.4	64.6	58.2	59.4	559,914
n	201,902	92,457	265,555	–	559,914

Geography of turnout

In Figure 5 earlier in this chapter we illustrated the negative relationship between ethnicity and turnout at the constituency level. However, we questioned whether this ecological relationship held at the individual level. The results detailed now indicate that it does not. To illustrate how this ecological fallacy arises, our sample was divided into separate categories according to the percentage South Asian living in the ward at the 2001 Census.[2] Four categories were chosen ranging from less than 5 per cent to wards where South Asians made up more than 20 per cent of the population.

Table 11 shows the percentage turnout by religion for these four categories. Quite clearly, overall, South Asian turnout increases where the South Asian population is more concentrated. The reverse is true for non-Asians. It seems that South Asians may live in areas of lower than average turnout, but this is precisely where they are most likely to vote.

Table 11 Percentage turnout (weighted) by religion and per cent South Asian in sample wards (design and vote weight – Vgweight)

Religion	0–4.9%	5–9.9%	10–19.9%	>20%	Total
Hindu	55.5	56.3	65.5	66.6	61.3
Muslim	56.1	54.6	60.7	61.5	58.5
Sikh	49.8	55.8	64.3	64.4	59.7
Other South Asian	55.8	54.3	63.9	57.2	55.8
All South Asian	55.1	55.7	62.2	63.4	59.4
Non-Asians	58.7	55.7	53.8	52.2	58.4
Total	58.7	55.7	55.2	56.3	58.3

Regarding the three main South Asian religious groups (Hindu, Muslim, Sikh), turnout tends to be higher where South Asian population is more concentrated. In wards where the South Asian population was more than 10 per cent, Sikh turnout was almost 15 percentage points above the equivalent rate for areas with less than 5 per cent South Asians. For Hindus the equivalent differential was over 10 per cent and for Muslims approximately 5 per cent. By contrast, non-South Asian electors were least likely to vote in areas with larger South Asian populations.

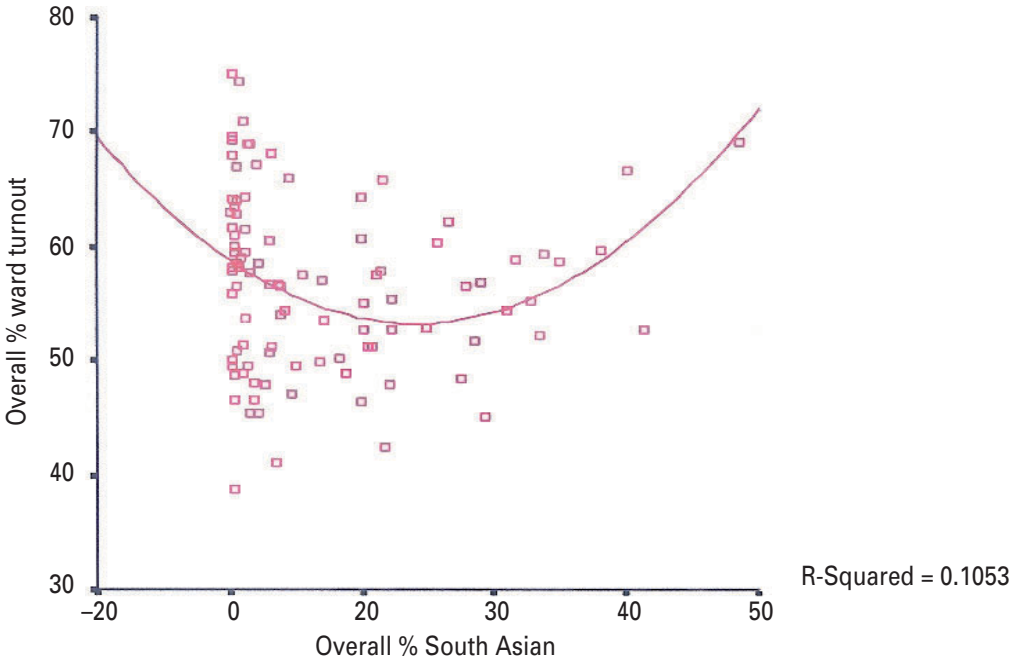
The results suggest that, while those of Indian heritage have been the most educationally and economically successful over recent years, and remain the people most likely to vote in general elections, the role of the extended family and strong community networks may still play a vital role in mobilising Hindu and Sikh voters. Interestingly, whereas it was Muslim communities that had the most notable ‘enclave’ effects in relation to registration, for turnout it was the Hindu and Sikh populations. It may be that, in areas with large Muslim populations, the trend towards higher registration rates is diluting otherwise higher turnout rates by encouraging more unlikely voters onto the register. The trend in turnout figures for non-Asian is the mirror image of South Asian turnout, clearly illustrating why the ecological relationship is misleading.

Are ecological analyses flawed?

Following from the above, if South Asians live in low turnout areas, ecological analyses would suggest that South Asian turnout is lower than it actually is. However, the individual-level evidence suggests this is an example of ecological fallacy. Although there are methods of ecological analysis that ameliorate this, the only reliable way to demonstrate this is by referring to the individual-level data as we have done here (King, 1997). Table 11 above provided evidence that this might arise because the geography of turnout of South Asian electors is the mirror image of that of other electors. We can now look at that claim in slightly more detail.

Earlier we looked at the constituency-level relationship using constituency results and (1991) Census data (Figure 5). We can now look at the results from our sample aggregated to ward (see Figure 6). The correlation is much weaker at the ward level because the ecological fallacy is ameliorated by adopting a smaller geographical unit. Also the relationship is non-linear, an increase in the South Asian population being associated with lower turnout at low percentages of South Asian population, but increasing as we move into very high concentrations of South Asian populations.

Figure 6 Nature of aggregate relationship at ward level (from sample 97 wards; percentage turnout by percentage South Asian)



However, we know from our individual analyses that South Asian turnout is the same or higher than non-Asian turnout. Furthermore, as Table 11 showed, South Asian turnout in the sample is actually higher in wards where South Asian population is higher, yet non-Asian turnout is much lower. In Figure 7, we disaggregate turnout by Asian/non-Asian and re-examine this ward-level relationship.

Figure 7 Comparing percentage turnout of South Asians with percentage turnout of non-Asians against overall percentage South Asian at ward level (from sample of 97 wards)

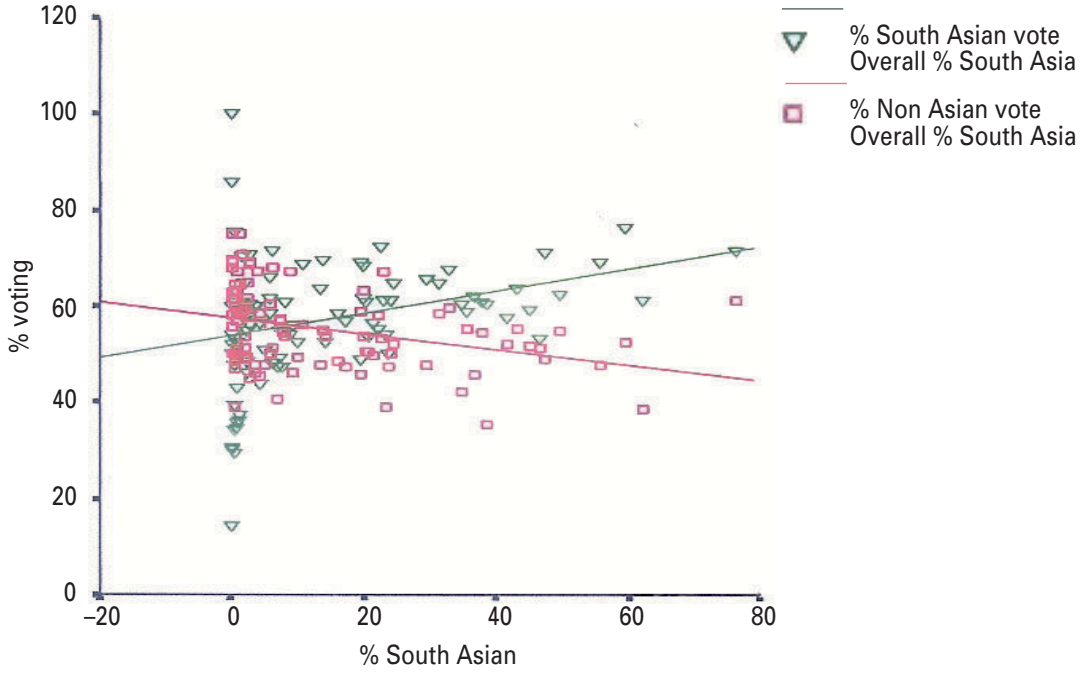
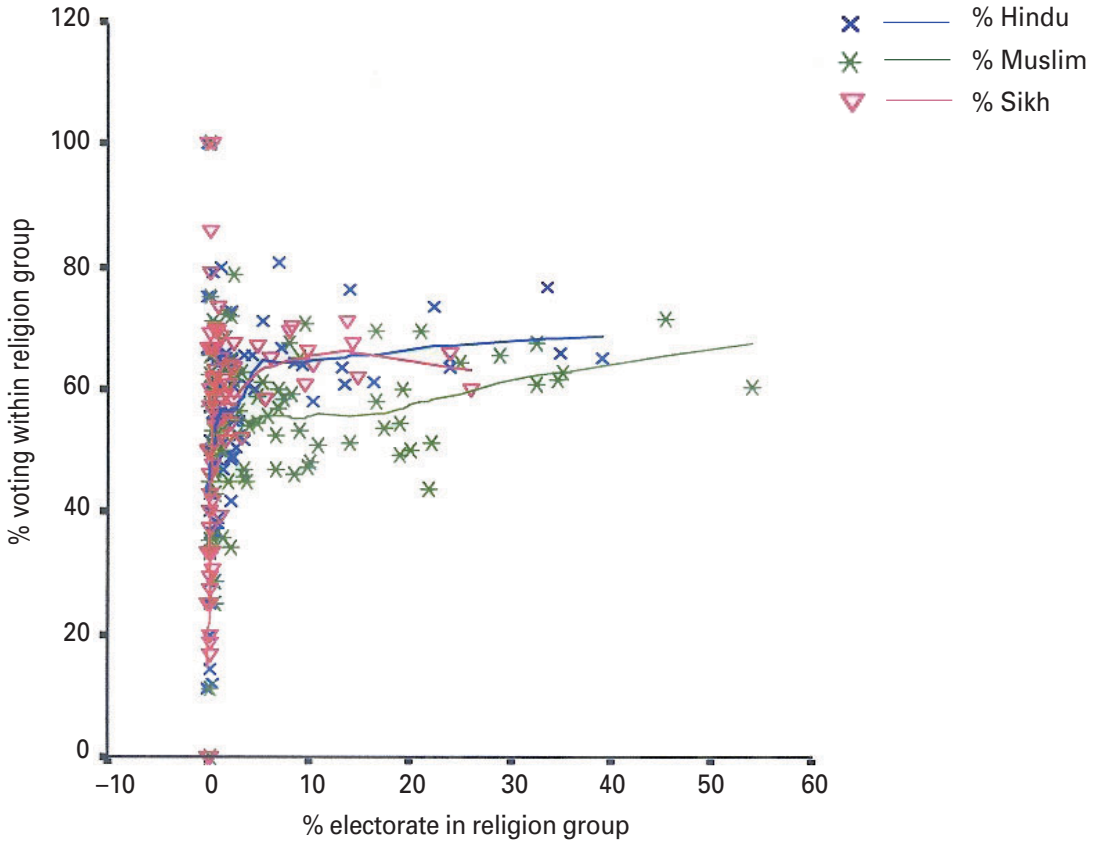


Figure 7 illustrates how the relationship between percentage South Asian electors and percentage turnout is positive for South Asian electors and negative for all other electors. This illustrates a classic ecological fallacy. For instance, wards such as University (Bradford), Charnwood (Leicester East), Whitefield (Pendle), Coldhurst (Oldham West and Royton), Limehouse (Poplar and Canning Town) contained 20 per cent or more South Asians and achieved South Asian turnout rates in excess of ten percentage points above non-Asian turnout. Yet, of the 38 sampled wards with a South Asian population of less than 2 per cent, only 12 recorded higher South Asian turnout rates than non-Asians. By contrast, only Headstone North (Harrow West), Costons (Ealing North) and Riverside (Cardiff West), of the 40 sampled wards with a South Asian population of more than 10 per cent, had a higher percentage of non-Asians voting than South Asian.

In Figure 8 we further disaggregate the South Asian turnout rates shown in Figure 7 into the three main subgroups. This shows that, for each group, but Muslims in particular, there was an increase in turnout as the proportion of the ward population made up by that group increased. These findings may make it difficult to rely on ecological results of BME voter turnout in the future.

Figure 8 Turnout by religious group against electorate share of that group



Registration and turnout

Traditionally, most western democracies, including France, which requires voters to take the initiative to register, have calculated electoral turnout in relation to the registered electorate. The one major exception has been the United States where the denominator in the calculation of official turnout is the 'voting-age population' derived from census data. According to US officials, using VAP to measure turnout is

more reliable, as it includes both those who failed to register and those who didn't vote. It is also less likely to overstate decreases in US turnout following large rises in registration brought about by recent changes to electoral law (1993 National Voter Registration Act). However, other problems associated with using the registered electorate as the denominator in the turnout ratio are not specific to the United States. For instance, the UK electoral register suffers from data-ageing problems as a result of new voters coming of age, people living at a temporary address or moving house and people who have died (Todd and Eldridge, 1987; Smith 1993). Similarly, the registered electorate can be inflated by people who are registered in more than one address – for example, students are often registered at both their home and term-time addresses. During the 16-month period between registers, 1.5 per cent of people die and 13 per cent of people move house (Pattie *et al.*, 1996). Smith (1993) estimated the level of redundancy on the register to be approximately between 1.8 and 3.3 per cent. The fallout from the Community Charge, particularly young men leaving the register (Smith and Mclean, 1994) and accusations that people deliberately avoid being on the register for a variety of reasons, e.g. to allow partner to qualify for reduced Council Tax.

Leading scholars, particularly from the United States, have used the VAP as the denominator in the calculation of turnout to substantiate the notion that electoral participation in modern democracies is in decline (Rosenstone and Hansen, 1993). Since Teixeira (1992) talked about the 'disappearing American voter', a number of empirical studies have reached similar conclusions (Wattenberg, 2004). Yet, using the VAP to measure turnout is not without its flaws. The VAP includes a large number of ineligibles (those without actual voting rights such as felons, people who do not meet residency requirements and non-citizens) and excludes eligible voters (military personnel and overseas electors). This has recently led to accusations from some that its inclusion in the turnout ratio masks real trends. For instance, McDonald and Popkin (2001) stress that, by redefining the denominator as the 'voting eligible population' (VEP), the presumed steady decline in US turnout barely exists (also see early work by Burnham, 1985, 1987 on estimating the number of eligible voters).[3] Aarts and Wessels (2002) also dismiss the VAP definition as 'a blurred measurement of turnout as the mobilising power of a system'. Since 2001, the VEP estimate derived by McDonald and Popkin is now widely used by US scholars, although a second, less sophisticated VEP estimate has also been put forward. This builds on earlier work by Burnham (1985, 1987) and simply calculates the turnout denominator as the VAP minus non-citizen adults. It has drawn controversy given that it estimates 2004 US turnout as exceptional, higher than in 1992, while both the other measures record 2004 turnout figures that are below 1992 estimates. The omission of eligible

expatriates from the second VEP estimate has drawn criticism, particularly given McDonald and Popkin's (2001) claims that the number of ineligible felons was much lower than overseas electors at all presidential elections from 1948–92. Althaus (2005) argues that Gans' estimates of turnout are therefore questionable, at least up to 1992 and probably beyond.

Yet, despite its flaws, using VAP in the calculation of turnout has become the worldwide benchmark (IDeA, 1999). Also, removing ineligible electors from the turnout denominator can be extremely difficult and imprecise (Teixeira, 1992). It is also important to stress that the US is an unusual case. For instance, it has a considerably higher number of felons and eligible expatriates than the UK.[4] While there are an estimated two million potential UK overseas electors (2003 figures), only an estimated 13,000 are actually registered. Registration laws also vary considerably among US states and between elections, making valid comparisons implausible.

2001 South Asian turnout after adjusting for registration

Bearing these considerations in mind we have recalculated the turnout rate of each group, having adjusted for the registration rates (with and without country of birth adjustments) calculated in the previous chapter. These are reported in Table 12.

Table 12 Turnout adjusting for registration (all results corrected for national turnout and excludes postal voters)

Religion	Adjusted % voted corrected for national turnout	Turnout after registration (with COB adjustment)	Turnout after registration (without COB adjustment)
Overall	58.3	58.5	57.4
Non-South Asian	58.3	58.6	57.7
South Asian non-Muslim	60.7	55.1	53.5
Muslim	58.5	53.5	48.0
All South Asian	59.4	55.5	51.2

Note: the all South Asian rate is higher (with COB adjustment) because of a larger number of valid OAs than both the South Asian subgroups (see Chapter 3 for more details).

Recalibrating the estimates of turnout to take into account different registration rates (without country of birth adjustments to registration) reveals that, as a percentage of the VAP, South Asian participation rates are lower than those of the rest of the population. This is because, while having high turnout among registered electors, South Asian communities have slightly lower levels of registration than non-South Asians (see Chapter 3). However, as we saw above, this is partly because of difference in eligibility arising from nationality. If we recalibrate turnout to registration rates adjusted for country of birth the difference is much smaller, with all South Asians having an overall turnout rate only 3 per cent lower than the population as a whole. As has been argued by other researchers (e.g. Wattenberg, 2004), lower registration rates can be linked to higher levels of turnout because those registered are more committed to voting. In other words those not registered would probably not vote anyway. Attempts to increase registration rates among South Asian voters might, therefore, pull South Asian turnout rates back towards the rate for the rest of the population.

Conclusion

This chapter describes what we consider to be the largest and most systematic nationally representative estimate of electoral turnout (free of response bias) among British South Asian communities ever undertaken. A number of important conclusions, which challenge orthodox perceptions, emerge from this unique study of South Asian voting. Our key findings are as follows.

- South Asian turnout among registered electors was higher than non-Asian in 2001.

- Even though South Asians tend to live in areas where there is lower than average turnout, it seems that they are more likely to participate in general elections than non-Asians.

- The figure of 59.4 per cent arguably represents the most accurate estimate of turnout among South Asian voters ever achieved, although, once these figures have taken account of voters who were not registered, overall participation rates are lower for South Asians, particularly Muslims.

- After adjusting for country of birth, or in other words estimating the turnout as a percentage of the voting eligible population, we find that there is very little difference between religious groups and, overall, the difference between South Asians and the rest of the population is only 1 per cent.

- South Asians of Indian heritage (Hindu and Sikh) have higher rates of participation than Muslims.

- Hindus were found to be the most active electors. Hindu turnout in 2001 was statistically significantly higher than the overall rate.

- Turnout was more than six percentage points higher among South Asian women than men, contradicting previous work based on survey data.

- Muslim women are more likely than non-Asian women to vote.

- While ecological analyses stressed the negative relationship between ethnicity and turnout, we demonstrated that the ecological relationship does not hold at the individual level.

- Using individual-level data, we reaffirmed this ecological fallacy by illustrating that South Asian turnout is highest where there are more South Asians in the electorate, which is where turnout for the rest of the population is lower.

- The strength of community networks, extended families and effective mobilisation are possible explanations for this pattern.

4 Factors affecting turnout

Key findings

This chapter attempts to understand some of the variation in South Asian turnout by using statistical models. We adopt a multilevel logistic regression model, which takes into account the clustered and stratified nature of our sample design. We find that the household is the most important unit of variation for turnout of all religious groups. Also South Asian voters of all religions were more likely to vote than their non-South Asian counterparts. And, for South Asian electors, the size of one's own religious group in the area was important in enhancing turnout, strengthening the hypothesis that South Asian communities are more effectively mobilised by political parties or community leaders.

Introduction

In the previous chapter we showed how South Asian turnout compared with that of the rest of the population. In particular, we showed that, among registered electors, turnout of South Asian electors was slightly higher than that of other electors, and the Hindu electorate had the highest rates. In many ways it is heartening that such small differences exist. In other ways it is somewhat surprising. First, it is surprising because it contradicts aggregate-level analyses, as we have seen, because turnout for South Asian electors is higher in areas where turnout of the rest of the population is lower. Second, South Asian communities have very different characteristics, which might be expected to affect relative turnout levels. Both these reasons relate to general factors affecting turnout since they relate to factors that potentially affect the whole electorate, albeit in different ways. General factors might be usefully divided into 'individual' effects such as age or social class, or housing tenure (Le Lohe, 1990) and 'systemic' effects relating to the operation of the electoral system (such as the difference between the parties, whether one lives in a marginal seat or the closeness of the election overall). These factors may not affect BME groups equally. A third reason these results may be surprising relates to community-specific factors. There are a variety of community-specific reasons why turnout might be expected to be lower in South Asian communities. These are likely to relate to issues of representation (e.g. failure to represent particular views or lack of BME candidates), mobilisation or lack of mobilisation, a diminished sense of effectiveness, social or economic exclusion, and racism (Geddes, 1998; Messina, 1998). The impact of these factors might be transmitted via measurably different attitudes towards (for

example) political institutions or politicians, or different levels of interest in politics or feelings of civic duty (Purdam *et al.*, 2002).

In this chapter we attempt to understand some of the variation in South Asian turnout by using statistical models. While such models are designed to uncover the correlates of turnout for different groups we do not seek to understand individual motivations. This requires a different research methodology, and is beyond the scope of this research.

Models of turnout

Statistical models can be effective tools for identifying and quantifying factors that affect various social outcomes. In brief, statistical models are generally used to test hypothesised relationships between different variables (often measuring people's characteristics) while holding constant other variables. For example, we might want to examine the relationship between gender and wages after controlling for (or holding constant) the number of hours worked or the occupations of men and women. The simplest statistical model for this kind of problem is ordinary least squares regression. However, because our outcome of interest is categorical (to vote or not to vote) rather than continuous (e.g. wages), a different model is required. The most common model for these kind of data is called logistic regression since this is suitable for a binary outcome. However, rather than using a simple logistic regression model, we adopt a rather more sophisticated, multilevel logistic regression model, which takes into account the fact that our sample is not a simple random sample but rather the design included clustering and stratification (see Chapter 2), and that people living close together tend to be relatively homogeneous (that they are more alike than people selected at random). An added advantage is that we can estimate the amount of variation at each of the levels in the model, and therefore observe (for example) the relative importance of household and neighbourhood influences where these are not measured by other variables in the model (e.g. whether or not abstention runs in households).

Using multilevel models

Multilevel models can have any number of 'levels' depending on the structure of the population being modelled. In our analyses we have already identified four different levels in which electors operate. These are the individual elector, the household in which he or she lives, the immediate neighbourhood as measured by the Census

output area and the primary sampling unit from which they were drawn (the ward).[1] One key obstacle to overcome in our analyses is a lack of individual data. Electors are individuals and have individual characteristics and attitudes. Unfortunately, the electoral registers from which our data are drawn contain no information about the electors other than their name. As already discussed extensively this has been used to derive the religious group of South Asian electors. However, we know nothing of their other characteristics, nor their attitudes or values. Thus, as noted above, we do not attempt to understand the individual motivations of electors for voting, which might be more effectively measured using qualitative approaches or survey data. We do, however, have information about the social situation of voters, which can be gleaned from the elector's address.

As mentioned earlier, people living together tend to be relatively homogeneous and the smaller the geographical areas the more homogeneous the population. Fortunately for us, census output areas, which we have identified for all electors in our sample, were designed explicitly to be homogeneous. We can therefore infer a lot about our sample from census data for output areas in conjunction with the address of each individual (notwithstanding the possibility of 'ecological fallacy'). Furthermore, because the Census provides a significant amount of information about ethnic and religious subgroups of the population, as well as data on the population as a whole, we are able to create some variables that specifically relate to the religious group in question. Such variables, when measured at the local level (e.g. the OA), can be used as proxies for individual data. For example, the percentage of Muslims who are owner-occupiers can be thought of as probability that a Muslim from that OA is an owner-occupier. In other words, although it is area-level data, we can interpret it as a measure of individual characteristics. In other instances, particularly where variables relate to wider populations or larger geographical areas, it makes more sense to treat the measures as contextual variables. Contextual variables can best be understood as variables that capture the impact of living in a particular type of area, regardless of one's personal characteristics. This might, for example, be the impact of living on a council estate over and above the impact of living in council accommodation.

Many variables in our model will capture elements of contextual as well as individual-level effects, but in Table 13 we have illustrated how we think different variables used in the model are best classified. Some variables in the model are purely contextual since they only measure area characteristics, not elector characteristics (e.g. ethnicity of candidates). It is important to note that variables are included at only one level and are specified as religion specific wherever the data allow. Lower levels (e.g OA) are tested before higher levels (e.g. ward) and the latter are only included where the former are not significant. A full list of variables is provided in Appendix 6.

Table 13 Levels and example variables in the models

Level	Personal/household characteristics	Religion-specific individual proxies	Generic individual proxies	Generic contextual
Individual	Religious group	NA	NA	NA
Household	Household size	NA	NA	NA
Output area	NA	% home ownership	% low social class	% Muslim
Ward	NA	% degree	NA	% non-migrant
Constituency	NA	NA	NA	Marginality

In summary, using information on ethnicity and turnout at the individual level in conjunction with contextual data about the areas in which people live, we examine the relative importance of the individual's religion and the characteristics of the area in which s/he lives. Information about the ethnicity of the candidate and local social/political context is measured at the ward/constituency level and related directly to the turnout of South Asian voters. The first step in our modelling strategy is to examine the extent of variation at the different levels. Then we proceed to attempt to ask whether this variation is accounted for by other variables in the model and also whether differences between religious subgroups persist after controlling for other factors. Having done this we approach the problem from a different angle and attempt to determine whether different factors affect different subgroups in different ways. This is achieved by separately modelling our subgroups.

Sources of variation

We know from experience and other research that turnout varies between neighbourhoods, between constituencies and also between individuals and households. Our analysis enables us to look at the extent of variation for different groups in the population at the different 'levels' in the model (Table 14). What is immediately apparent is that, in all models, household variation is relatively large compared to variation at other higher levels, suggesting that 'people who live together vote together'(see also Johnston *et al.*, 2005).[2] However, household variation is lower for Hindus than for other groups, especially non-Asians. Variation at the level of the output area, ward and constituency was rather small in comparison in all models, and indeed in some instances was zero and was therefore not included in the model. In general there is more variation at the level of the OA than the ward, reflecting the fact that people who live close together tend to be more similar. However, ward and constituency variations tend to compete with each other, as in many cases they are one and the same. Combined, ward and constituency variations usually exceed OA variation, suggesting that there is something about these higher-level political units that is important, above and beyond the fact that

they contain similar electors. One explanation of this may be that turnout varies between constituencies (and hence wards in our sample) because of the different levels of competitiveness and campaigning between constituencies. Constituency variation is rather higher for Hindus than for the rest of the population suggesting that the constituency location for Hindus is relatively important.

Table 14 Variance estimates for variance components models

	Household	Output area	Ward	Constituency
All people*	0.879	0.088	0.047	0.037
Muslim	0.857	0.055	0.065	0.030
Hindu	0.716	0.045	–	0.123
Sikh	0.849	0.065	–	0.054
Non-Asian*	0.836	0.095	0.109	0.001

* *10 per cent sample of households.*

When a number of likely explanatory variables are added to the model (see below) we would expect that this would account for some of the variation in turnout at the higher levels (OA, ward and constituency). The same would not be true for households since we have no household-level control variables. Figure 9 shows the percentage of the original variance that remains at each higher level after controlling for a range of explanatory variables (detailed below). The figure shows that social and political characteristics of the areas reduce the ward variation very considerably where such variation existed (i.e. for non-Asians and Muslims). The models also explain a substantial proportion of variation in turnout between OAs, particularly for non-Asians and Muslims, but less so for Hindus and Sikhs. However, for the latter, the models explain nearly all the constituency variation and, for the Muslims, all constituency-level variation is accounted for. In other words there are localised variations in South Asian turnout that cannot be accounted for by social and political factors, but, once these factors have been taken into account, constituency variation is almost non-existent.

Differences by religion and influences on turnout

The first set of models we looked at included all electors in sampled wards, excluding postal voters and non-eligible voters as discussed in the previous chapter. One of the main aims of this approach was to look at the extent to which religion differences could be accounted for by social and demographic variables discussed above. Table 15 gives the coefficients for each religious group before and after controlling for all other significant explanatory variables. The other significant predictors of turnout in this model (i.e. for all groups combined) are listed in Table 16.

Figure 9 Percentage of variance in null model remaining in full model by geographic level and religion.

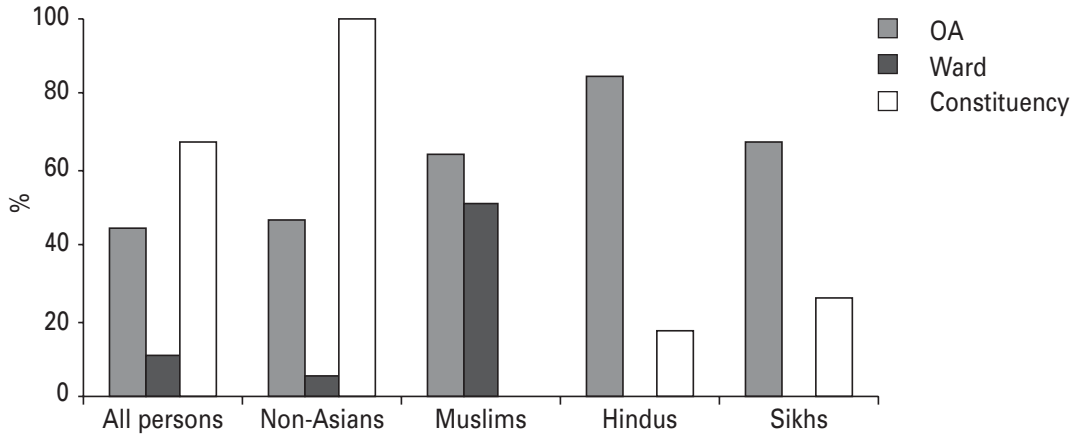


Table 15 Coefficients for religion, model of all electors

Religion	Model 1 (religion only)	Full model*	% religious differential explained
Hindu	+0.417	+0.380	8.9
Sikh	+0.358	+0.308	13.9
Muslim	+0.323	+0.277	14.2

* *Including all variables listed in Appendix 6. Reference group is non-South Asian.*
Note: based on a stratified sample of 10 per cent of households. All coefficients are statistically significant (p = 0.05).

Table 16 Additional significant effects, all electors

Variable	Effect
OA variables (% all adults)	
Homeownership	+0.007
Manufacturing	+0.005
Degrees	
Pensioners	+0.005
Unemployed	–0.023
Muslims	+0.007
Low social class	–0.009
Managerial	+0.004
Constituency	
Margin	–0.006

The results for the null model (which allows for the sample design) in Table 15 confirm that registered Hindus are the most likely group to vote, followed by Muslims and Sikhs, all more likely than the rest of the population. Even after controlling for other variables there are significant religion differences and the coefficients are only slightly reduced by controlling for the social characteristics of the areas in which different groups live. It therefore appears that the differences we observed in turnout are not the products of social structure but genuine differences in the propensity to vote or other unmeasured effects.

The differences by religion reported in the third column of Table 15 are the net religion effects after taking into account all the other variables in our models. Those that are significant are reported in Table 16. We see that a number of social characteristics of OAs impact on turnout, presumably because of the different propensity of these groups to vote. First, there is evidence of increased propensity of older voters to turn out; the percentage of pensioners in the OA is strongly positively associated with higher levels of turnout. This is consistent with individual-level analyses that show a similar pattern. Sociological and resource-based theories of turnout hypothesise that those with more resources and greater social status are more likely to turn out to vote, and similarly (according to relative deprivation explanations) the more socially and economically deprived are less likely to vote. In keeping with this, we find that turnout in our sample is positively correlated with homeownership and the percentage in managerial and professional occupations, both of which are good indicators of high social status or affluence. Conversely, the greater the percentage of the OA population in lower social classes and the greater the number of unemployed, the lower the level of turnout. However, contrary to expectations, the percentage in manufacturing industries is positively associated with turnout and the percentage with degrees is negatively related. Both are likely to be correlated with other social characteristics mentioned above, so may simply offset other class effects.

The percentage of the population in the OA who are Muslim is positively related to turnout, even after controlling for individual-level religion. This indicates that, given the types of social area in which Muslim populations live, the turnout in those areas is higher than might be expected from the social profile alone.

We also find that, the more marginal the constituency in which the individual lives, the more likely people are to vote. This would be expected if electors were rationally responding to the chance of affecting the outcome of the election, or parties were mobilising voters more effectively where they were most needed (and is consistent with much other research; see Denver and Hands, 1997; Johnston and Pattie, 1998).

Different religions, different factors?

Above we saw that Hindus were the group most likely to vote, but what factors affect turnout for the Hindu population? Similarly, Muslims had lower levels of turnout than Hindus and Sikhs, and this could only partially be accounted for by social and demographic differences. We have already seen that the turnout of each group varies with the proportion of electors that group comprises in the ward. Using the statistical models described above, but applying them to each of the religious groups in turn, we can observe which other factors affect turnout for each group. Table 17 provides a breakdown of the factors that were significant in explaining variation in turnout of each religious group.

Table 17 Significant effects on turnout by religious group

Variable	Hindu	Sikh	Muslim	Non-Asian
OA variables (all persons)				
Hindu %	+0.008	+0.008		
Muslim %		+0.006	+0.007	
Sikh %		+0.012		
Degrees %	–0.006		–0.008	
OA variables (religion specific)				
Homeownership %	+0.003	+0.003	+0.003	+0.009
Pensioners %	+0.002		+0.002	+0.009
Students %				
Low social class %	+0.013		+0.019	
Long-term illness %				–0.008
Ward variables (all persons)				
Students %	–0.008			
Same address %				–0.008
Ward variables (religion specific)				
Degrees %				+0.008
Two cars %				+0.011
Constituency				
Margin 97%	–0.005	–0.008	–0.007	–0.002

The results of the models summarised in Table 17 reveal a number of common underlying factors that influence South Asian turnout. Indeed, given the number of variables that were tested, there is a remarkable degree of communality between factors affecting turnout in different South Asian communities, yet at the same time a number of different factors seem important for the rest of the population. Dealing first with the religion-specific variables, homeownership at the OA level was significant for all religious groups, and was positively associated with turnout. The percentage of pensioners in the religious group was also positively associated with turnout for Hindus, Muslims and non-South Asians, but not for Sikhs (though the variable was only marginally insignificant after including other variables in the model). Hindu and

Muslim turnout is also influenced by the social class status of the Hindus and Muslims in the output area and the ward respectively. Hindu turnout is also lower in wards with a large number of students. Hindus and Muslims have positive effects for the lower social classes (routine and semi-routine occupations) in the OA. This appears to contradict the general pattern of higher social classes being more likely to vote. However, this is likely to reflect a complex relationship between social structure and the level of participation and the ethnic composition of neighbourhoods. In short, areas with concentrations of the lowest social classes tend to be urban and metropolitan, which is exactly where South Asian turnout is highest. In contrast, turnout for both these communities is lower where there are a larger number of degree holders among the population as a whole, which generally tends to be in more middle-class areas.

This interpretation is supported by positive effects for the proportion of the population made up by Muslims (significant for Sikhs as well as for Muslims), Hindus (significant for Hindus and Sikhs) and Sikhs (significant for Sikhs only). This not only confirms the mobilising effect of living within religiously diverse areas where minority electors live within sizeable communities, but also suggests this effect is not always restricted to those living in the particular religious community to which they belong. For non-South Asian electors, the religious profile is not significantly related to turnout. However, for this group, a number of variables are significant that were not so for the South Asian groups. In particular population stability suppresses turnout (contrary to expectations), as does the proportion of people with a long-term illness, while manufacturing, the highly qualified and multiple car ownership enhance turnout.

At the constituency level, the marginality of the seat proved to be important for all groups, the size of the margin having a negative relationship with turnout. This is consistent with what would be expected if electors are responding to the competitiveness of the race.

Conclusion

This chapter uses multilevel logistic regression models to understand some of the variation in South Asian turnout. A number of important findings emerge from the multilevel analyses.

- The household is the most important unit of variation for turnout of all religious groups. In other words people who live together are more likely to vote together. However, this effect was slightly weaker for Hindus than other groups.

- There is a small (but significant) amount of variation between output areas and constituencies in the propensity to vote of each group.

- Constituency effects were largest for Hindus.

- South Asian voters of all religions were more likely to vote than their non-South Asian counterparts, even after controlling for the clustering in the sample and the characteristics of the areas in which they live.

- Only a very small percentage of the religion differential could be accounted for by social differences.

- For the results of separate models of each religion we found different factors being influential than for the non-South Asian population. The only common factors in all these models were homeownership, which was positively associated with turnout of all groups, and the degree of marginality of the constituency.

- For South Asian electors, the size of one's own religious group in the area was important in enhancing turnout, lending further support to the hypothesis that South Asian communities are more effectively mobilised by political parties or community leaders.

- These findings are consistent with social capital theory, which suggests that more socially connected communities are likely to have higher levels of participation.

- The effect is sufficiently strong that it is in areas where South Asians are more likely to be from lower social classes where their turnout is highest, rather than in areas where more middle-class South Asians live, which tend to be in less ethnically diverse neighbourhoods.

5 Conclusions

This research has used the 2001 Census together with information from marked electoral registers from the 2001 General Election to provide a unique analysis of electoral turnout and registration among Britain's South Asian communities. There are a number of advantages to our methods. First, we have analysed actual rather than reported voting behaviour, thus removing the widespread problem of reporting on non-response bias that survey researchers have experienced. Second, we have analysed a sufficiently large sample to allow detailed analysis of subgroups in the South Asian population. Third, we have included estimates of registration on our research in order that we can calculate voter participation after adjusting for different levels of voter registration. This allows us to make the first comprehensive and reliable, nationally representative estimates of South Asian electoral participation in Britain.

Despite common perceptions that minority ethnic electors are less likely to vote in general elections than other electors, we have provided evidence that registered South Asian electors are actually more likely to turn out to vote. However, this is tempered by the finding that South Asian adults are less likely to be registered to vote than the rest of the population. Furthermore, like McDonald and Popkin (2001), we found that, by calculating turnout as a percentage of the voting-age population without adjusting for ineligibility due to birthplace, we underestimate turnout. After adjusting for country of birth, or in other words estimating the turnout as a percentage of the voting eligible population, we find that there is very little difference between religious groups and, overall, the difference between South Asians and the rest of the population is only 1 per cent.

While lower registration among South Asians, especially Muslims, is partly attributable to a larger proportion of the population being born outside of eligible countries, the fact that the turnout rate after adjusting for registration is lower than the rate for non-South Asians suggests that focusing on the reasons for non-registration may be equally as important as tackling non-voting. However, reasons for non-registration are currently poorly understood. Our models show that, in addition to the proportion born outside of the UK, Europe and the Commonwealth, other factors that affect South Asian registration include the level of unemployment, the number of older people and the extent of homeownership in the local area.

Muslim adults are more likely to be registered in areas with larger Muslim populations, but the equivalent pattern is not so clear for other South Asian adults. This 'enclave effect' is also apparent in patterns of turnout, and the relative size of

the local religious communities proves to be significant in multivariate models of Hindu and Sikh turnout as well as Muslim turnout. It is likely that political parties and community leaders play an important part in mobilising South Asian voters, especially Muslim voters, in terms of persuading them both to register and to vote. The tendency for higher turnout among South Asian electors in areas with larger South Asian populations, coupled with lower levels of turnout in those same areas, means that aggregate-level analyses can be misleading. In other words, while there is a negative correlation between the size of the South Asian population and turnout, this does not indicate low turnout of these groups, merely low turnout of their neighbours. If anything, South Asian electors are considerably boosting registration and turnout in inner-city areas.

Like registration, turnout varies between religious groups. Hindu electors are the most likely to vote of all the identifiable religious groups common in the South Asian electorate. Sikh turnout is also relatively high, while Muslim turnout is very close to the overall mean. However, contrary to previous research, we have found that Muslim women are considerably more likely to vote than Muslim men and nearly as likely as Hindu women. Turnout was more than six percentage points higher among South Asian women than their male counterparts. As already noted, all the identifiable South Asian groups turn out in greater proportions in areas where they are most concentrated. This may be a result of enhanced mobilisation effects in more diverse areas.

In Chapter 4 we showed that the household is the most important unit of variation for turnout of all religious groups, though this effect was slightly weaker for Hindus than for other groups. There is also a small (but significant) amount of variation between output areas and constituencies in the propensity to vote. The models confirmed that, after allowing for the clustering in the sample and the characteristics of the areas in which they live, South Asian voters of all religions were more likely to vote than their non-South Asian counterparts. Only a very small percentage of the religion differential could be accounted for by social differences. In part this reflects the fact that different factors are influential than for the non-South Asian population; using separate models for each ethnic group we found that the only common factors in all these models were homeownership (which was positively associated with turnout of all groups) and the degree of marginality of the constituency. The models also confirmed that South Asian electors are more likely to vote where they are most geographically concentrated. This provides support for the hypothesis that South Asian communities are more effectively mobilised by political parties or community leaders than other electors and is consistent with social capital theory, which suggests that more socially connected communities are likely to have higher levels of participation.

Notes

Chapter 1

1 http://www.electoralcommission.gov.uk/.

2 This is an example of the ecological fallacy (see Robinson, 1950).

3 It is not possible to derive ethnic origin of black Caribbean voters from a names analysis, thus, although the analysis by ethnicity is partial, this does not detract from the importance of understanding the participation of South Asian communities.

4 A similar approach, but not differentiating by religion, was used by Swaddle and Heath (1989).

Chapter 2

1 Because of the possibility of dual and acquired citizenship, ONS assumed that people born in countries outside of Europe and the Commonwealth were eligible to be registered (see earlier in Chapter 2).

2 Dorling (2007, forthcoming) attempts to validate the Census against the electoral register in Britain for county councils or local authorities.

3 Dorling (2007, forthcoming) uses mid-year estimates and points out that data correction to populations in the City of London, Westminster, etc. following lobbying by these councils has only increased the discrepancies.

4 The VAP in each ward includes over 18s and 10.68 per cent of 17 year olds (39/365 = 10.68 per cent).

5 Ninety-seven per cent of electors were successfully allocated a postcode and 1,823 out of 3,192 OAs were retained as valid under the criteria described. This is described in more detail below.

6 Empirical analyses show that cells with counts of 6 and under are affected by rounding. Because, here, adjustments are made for deaths and attainers, cells slightly greater than 6 were also affected. We therefore employed a cut-off midway between 6 and 7.

7 The apparent heteroscedasticity in Figure 3 is alleviated in the regressions, as the regressions are weighted to reflect the numbers of electors in each observation (which, for South Asian electors, are very unevenly distributed). Effectively, the extreme values around the zero value on the x-axis have tiny regression weights.

8 Diagnostic statistics revealed a small number of influential cases (standardised residuals greater than 3), which turned out to be output areas with very small denominators. There is no evidence of multicollinearity in any of the models. Variance-inflation factors were well within the established criteria for all predictors.

9 The positive coefficient for students is perhaps surprising, but should be treated with caution. Students should be recorded in the Census at their term-time address and may be registered to vote at both term-time address and address outside of term time where different. It is unclear as to the extent to which the Census instructions were followed regarding students living away from home and the extent to which students register at either or both addresses. It should also be noted that the bivariate correlation for the full-time student variable is small and negative, suggesting the positive model coefficient is affected by correlation with other explanatory variables.

Chapter 3

1 The total percentage turnout and non-Asian turnout are similar – if we went more than one decimal place you would find that the total would be a little higher. Basically, the South Asian aspect is so small, it has only a small impact on the total.

2 Information obtained from 2001 Census data.

3 The VEP (voting eligible population) removes non-citizens, persons ineligible because of criminality and adds civilian and military personnel living overseas. This measure does not remove the number of mentally incompetent persons or people ineligible because of state residency requirements. The numerator is the total number of votes cast for highest office (e.g. in presidential years this is the number of persons who voted for presidential candidates).

4 The US prison population rate is 686 per 100,000 (2003 figures) compared to the UK prison population rate of 139 per 100,000 (highest among EU countries).

Chapter 4

1 The ward and constituency are treated as the same level, as in most constituencies we have sampled only one ward, making it impossible to distinguish ward from constituency effects.

2 Individual-level variation is not modelled, as it is constrained to have a binomial distribution with a variance of 1.

References

Aarts, K. and Wessels, B. (2002) 'Electoral turnout in Western European democracies', American Political Science Association Conference, Boston, MA

Ali, A. and Percival, G. (1993) *Race and Representation: Ethnic Minorities and the 1992 Elections*. London: Commission for Racial Equality

Althaus, S. (2005) 'How exceptional was turnout in 2004?', *International Communication Association & American Political Science Association*, Vol. 15, No. 1, http://www.ou.edu/policom/1501_2005_winter/commentary.htm

Anwar, M. (1990) 'Ethnic minorities and the electoral process: some recent developments', in H. Goulbourne (ed.) *Black Politics in Britain*. Aldershot: Gower

Anwar, M. (1994) *Race and Elections*. Coventry: CRER

Anwar, M. (1996) *Race and Elections*. London: Routledge

Anwar, M. (1998) *Ethnic Minorities and the British Electoral System*. Warwick: CRER and OBV, University of Warwick

Burnham, W. (1985) 'Those high nineteenth-century American voting turnouts: fact or fiction', *The Journal of Interdisciplinary History*, Vol. 26, No. 4, pp. 613–44

Burnham, W. (1987) 'The turnout problem', in J. Reichley (ed.) *Elections American Style*. Washington, DC: Brookings Institution

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

Cummins, C., Winter, H., Cheng, K.K., Maric, R., Silcocks, P. and Varghese, C. (1999): 'An assessment of the Nam Pehchan computer program for the identification of names of South Asian ethnic origin, *Journal of Public Health Medicine*, Vol. 21, No. 4, pp. 401—6

Denver, D. and Hands, G. (1997) 'Turnout', in P. Norris and N. Gavin (eds) *Britain Votes 1997*. Oxford: Oxford University Press

Dorling, D. (2007, forthcoming) 'How many of us are there and where are we? A simple independent validation of the 2001 Census and its revisions', *Environment and Planning A*

Electoral Commission (2001) *Election 2001*. London: Politicos

Electoral Commission (2005) *Election 2005: Turnout Black and Minority Ethnic Survey*. London: Politicos

Geddes, A. (1998) 'Inequality, political opportunity and ethnic minority parliamentary candidacy', in S. Saggar (ed.) *Race and British Electoral Politics*. London: UCL Press

Heady, P., Bruce, S., Freeth, S. and Smith, S. (1996) 'The coverage of the electoral register', in I. McLean and D. Butler (eds) *Fixing the Boundaries: Defining and Redefining Single-member Electoral Districts*. Aldershot: Dartmouth

Heath, A. and Taylor, B. (1999) 'New sources of abstention', in G. Evans and P. Norris *Critical Elections: British Parties and Voters in Long Term Perspective*. London: Sage

IDeA (1997) *Voter Turnout from 1945 to 1997: A Global Report on Political Participation*. Stockholm: International IDeA

IDeA (1999) *Voter Turnout from 1945 to Date: A Global Report on Political Participation*. http://www.idea.int/Voter_turnout/parliamentary.html

IDeA (2002) *Voter Turnout since 1945: A Global Report*. http//www.idea.int/vt/survey/voter_turnout8.cfm

Johnston, R.J. and Pattie, C.J. (1997) 'Towards an understanding of turnout at general elections: voluntary and involuntary abstentions in 1992', *Parliamentary Affairs*, Vol. 50, pp. 280–91

Johnston, R.J. and Pattie, C.J. (1998) 'Voter turnout and constituency marginality: geography and rational choice', *Area*, Vol. 30, No. 1, pp. 38–48

Johnston, R., Jones, K., Propper, C., Sarker, R., Burgess, S. and Bolster, A. (2005) 'A missing level in the analyses of British voting behaviour: the household as context as shown by analyses of a 1992–1997 longitudinal survey', *Electoral Studies*, Vol. 24, pp. 201–25

King, G. (1997) *A Solution to the Ecological Inference Problem: Reconstructing Individual Behavior from Aggregate Data*. Princeton, NJ: Princeton University Press

Le Lohe, M.J. (1990) 'The Asian vote in a northern city', in H. Goldbourne (ed.) *Black Politics in Britain*. Aldershot: Gower

LGA (Local Government Association) (2000) *The Only Way is Up: Increasing Turnout in Local Government Elections*. London: LGA

McDonald, M. and Popkin, S. (2001) 'The myth of the vanishing voter', *American Political Science Review*, Vol. 24, No. 4, pp. 963–74

Martin, D. (2002) 'Output areas for 2001', in P. Rees, D. Martin and P. Williamson (eds) *The Census Data System*. Chichester: John Wiley & Sons

Mason, S., Gambles-Hussain, M., Leese, B., Atkin, K. and Brown, J. (2003) 'Representation of South Asian people in randomised clinical trials: analysis of trials' data', *British Medical Journal*, Vol. 326, No. 7401, pp. 1244–5

Mény, Y. (2002) 'France: the institutionalisation of leadership', in J. Colomer *Political Institutions in Europe*. London: Routledge

Messina, A. (1998) 'Ethnic minorities and the British party system in the 1990s and beyond', in S. Saggar (ed.) *Race and British Electoral Politics*. London: UCL Press

Nanchchal, K., Mangtani, P., Alston, M. and Dos Santos Silva, I. (2001) 'Development and validation of a computerized South Asian Names and Group Recognition Algorithm (SANGRA) for use in British health-related studies', *Journal of Public Health Medicine*, Vol. 23, No. 4, pp. 278–85

Norris, P., Lovenduski, J. and Campbell, R. (2004) *Gender and Political Participation*. London: Electoral Commission

Pattie, C., Dorling, D., Johnston, R. and Rossiter, D. (1996) 'Electoral registration, population mobility and the democratic franchise: the geography of postal votes, overseas votes and missing votes in Great Britain', *International Journal of Population Geography*, Vol. 2, pp. 239–59

Purdam, K., Fieldhouse, E., Russell, A. and Kalra, V. (2002) *Voter Engagement among Black and Minority Ethnic Communities*. London: Electoral Commission

Rees, P., Parsons, J. and Norman, P. (2005) 'Making an estimate of the number of people and households for output areas in the 2001 Census', *Population Trends*, Vol. 122, Winter, pp. 27–34

Robinson, W.S. (1950) 'Ecological correlations and the behaviour of individuals', *American Sociological Review*, Vol. 15, pp. 351–7

Rosenstone, S.J. and Hansen, J.M. (1993) *Mobilisation, Participation and Democracy in America*. New York: Macmillan

Saggar, S. (1998) *The General Election 1997: Ethnic Minorities and Electoral Politics*. London: Commission for Racial Equality

Smith, J. and McLean, I. (1994) 'The poll tax and the electoral register', in A. Heath, R. Jowell and J. Curtice (eds) *Labour's Last Chance? The 1992 Election and Beyond*. Aldershot: Dartmouth

Smith, S. (1993) *Electoral Registration in 1991*. London: OPCS Social Survey Division, HMSO

Steele, D. (2005) 'Identifying ethnicity: a comparison of two computer programmes designed to identify names of South Asian ethnic origin', UKPHA Conference, Gateshead

Swaddle, K. and Heath, A. (1989) 'Official and reported turnout in the British general election of 1987', *British Journal of Political Science*, Vol. 19, pp. 537–51

Teixeira, R. (1992) *The Disappearing American Voter*. Washington, DC: Brookings Institution

Todd, J.E. and Butcher, B. (1981) *Electoral Registration in 1981*. London: HMSO for OPCS

Todd, J.E. and Eldridge, J. (1987) *Improving Electoral Registration*. London: HMSO for OPCS

Wattenberg, M. (2004) 'The decline of party mobilisation', in R. Dalton and M. Wattenberg (eds) *Parties without Partisans: Political Change in Advanced Industrial Democracies*. Oxford: Oxford University Press

Introduction to appendices

The following appendices detail the methods used to calculate registration and turnout. The technical report in Appendices 1 and 2 includes a discussion of the sample design, accuracy of the marked registers, in-depth review of the names recognition software (Nam Pehchan and SANGRA) and a full account of the procedure taken. Appendices 3 and 4 focus on the methods used to create registration and turnout rates. For registration, this includes a detailed account of how we matched postcodes to census output areas and the procedures taken to create the numerator and denominator. For turnout, we review the use of weighting, including post-stratification weighting and our sample of postal votes. The final two appendices include details of the variables used in the statistical models.

Appendix 1: Technical report – sample and electoral registers

Sample and electoral registers

At all general elections, the electoral registers are marked manually according to whether each registered voter actually voted.[1] This research uses marked registers from the 2001 General Election for a representative sample of 97 wards, based on a stratified random sample (see Table A1.1). Using 1991 Census data, we stratified percentage South Asian using wards as the primary sampling unit.[2] Wards were sampled disproportionately in areas with a large Asian population to ensure the effective coverage of different subgroups, but weights are applied to make the sample nationally representative. All electors were included in the primary sampling units (see Table A1.1).

Table A1.1 Stratified random sample

	% South Asian population	No. of wards	Sample	Total population	South Asian population	No. of sampled electors (Census 2001)	No. of registered South Asian electors in sample
Stratum 0	0	2,057	0	NA	0	0	0
Stratum 1	>0–<0.5	5,134	18	140,030	314	99,068	534
Stratum 2	0.5–<2	1,972	20	125,955	1,495	88,509	1,615
Stratum 3	2–<10	1,025	19	158,849	6,721	116,310	7,128
Stratum 4	10–<20	201	20	187,869	27,669	131,905	24,827
Stratum 5	20+	163	20	225,984	85,372	155,348	65,197
Total		10,552	97	838,687	121,571	591,140	99,301

Note: all registered electors in the sampled wards are included, though this will be somewhat less than the total population (column 5), which includes all persons of all ages. The original sample was 100 wards, but because of data problems three wards were excluded. The number of sampled electors includes the 341 electors that were assigned a missing value in the dataset because information was lost due to photocopying error.

Accuracy of the marked electoral registers

The marked electoral registers list names by constituency, ward and polling district, and include the individual's title, forename, surname, address and postcode. A series of letters are used to indicate if the person can vote only in local or European elections. The registers also indicate if the person is a first-time voter and some contain the person's date of birth.

At an election, if a person turns up to vote at a polling station, a manual mark is made by the person's name by the recorder at a polling station. The register also includes whether the person has requested a postal vote. But whether or not the person actually voted with their postal vote is not recorded.

Across the 97 registers we found a wide range of differences and inconsistencies. There was considerable variation in how a vote was indicated. Some had a tick, others had a circle and some had a straight line through. In many instances this made it very difficult to assess whether someone had actually voted. A number of registers also included additional names entered by hand. Below we summarise some of the key findings.

Table A1.2 indicates the letters used on 81 of the registers and the definitions regarding elector eligibility. Despite this general uniformity on the classification of eligibility, the registers did vary widely in other areas. These inconsistencies meant that three distinct versions of the register could be identified, while a further eight registers contained a combination of all three features.

Table A1.2 Examination of inconsistencies and differences between registers

Electors' eligibility	Letters	Version 1	Version 2
Can vote only at local elections	G	✔	✔
Can vote only at European elections	E	✔	✔
Cannot vote at parliamentary elections	L	✔	✔
National of the EU and cannot vote at parliamentary elections	K	✔	✔
National of the EU and can vote only at European elections	U	✔	✔
Can vote only at parliamentary and European elections	F	✔	✔
Postal voter/cannot vote in person on polling day	A	✔	✔
Proxy	P	✔	✔
Classification of first-time voter	Date	✔	✘
Classification of first-time voter	x/line date	✘	✔
Removal of electors (deceased/moved, etc.)	Name removed or deleted entry	✔	v

The two versions account for 81 of the 97 registers. The most common format (version 1: 43 registers) included the standard letters outlined in Table A1.2 plus A and a line through the elector's name to indicate whether he/she was a postal voter. Version 2 (35 registers) included the letter A in the register but provided no letter or definition on the front sheet. There were also differences in the deletion of names. The former either used the words 'name removed' or simply crossed out the name with a double line. This made it difficult to ascertain whether an elector had voted or not, particularly when lines through names were used to indicate whether that person had turned out to vote. Version 2 used the words 'deleted elector' and double lines, sometimes within the same register, for the same purpose.

Differences were also apparent between the two versions in terms of the classification of electors reaching voting age. In version 1, registers include a date to show when the elector reaches voting age and then put a line through the elector to indicate that they could not vote in the 2001 Election because it preceded the date of eligibility (18th birthday). Consequently, no first-time voters before the 2001 General Election (7 June 2001) are indicated. By contrast, version 2 indicates those electors who are reaching the voting age within the life of the register. Either a line is drawn through the name to indicate that the elector cannot vote or the letter X is placed before the name. However, it is possible to determine the voting patterns for these first-time voters. Some registers don't place a line through the elector's name (underlined) or mark 'X' against the name, but the date given indicates whether he/she is eligible or not.

Eleven registers ignore the standardised format and use a more simplified version. Apart from the use of the letter A to define postal voters, a letter D is printed before the name to indicate that the elector is deceased and N to demonstrate that the elector is ineligible to vote at the parliamentary election. Regarding electors reaching voting age, U or Y is included before the name to indicate that the elector has not reached the voting age by polling day and cannot vote. Where a date exists and these letters don't appear, the elector is eligible to vote. In a similar manner to version 2, this allows the researcher to determine voting patterns of those electors reaching the voting age within the life of the register. However two inconsistencies exist. Among the 11 registers, the letter U was included on the front sheet instead of the letter Y, yet Y was used repeatedly throughout the register. Furthermore, on version 1 and 2 of the register, the letter U refers to those electors who are registered to vote only at European elections and has nothing to do with the eligibility of electors reaching voting age.

Of the remaining eight registers from the sample, one defined the letter A as absent voter despite including the standard format of letters. This seemed to refer to postal voters. Another register used variations of both version 1 and 2 to define postal voters – some electors have letter A and a line through their name, while others simply have a letter A. It is therefore impossible to know whether the register actually recorded those postal voters who had voted. The remaining six registers didn't include a front sheet with definitions and additional information. Again there seem to be inconsistencies, with some postal voters marked as voting in one or two polling districts and not in others. There was also no evidence of deleted entries, although double lines were used to indicate whether the elector had voted or not, thus making it awkward to differentiate. Also there were a number of registers that included marks indicating that the elector had voted where letters such as K or G existed before the name to show that the elector was ineligible to vote. In some cases, the letters CC

(county council) were written on the register to highlight the ineligibility of the elector at the 2001 General Election.

Notwithstanding these variations between registers, other inconsistencies were found. Given that local elections took place in some areas on the same day as the General Election, a number of marked registers included marks for both. Often, + and – or X is used as a mark to show voting. However, as stated earlier, X was used in the classification of first-time voters on some registers. The use of such marks is extremely confusing, as it is very difficult to decipher whether the + and – or X refers to the elector voting in the parliamentary or local election. Fortunately, in some registers, the inclusion of the letter G (can only vote in local elections) made it possible to differentiate between the two elections. Otherwise, one assumed that turnout would be greater at the parliamentary rather than at the local level.

Several registers include the words 'electors not registered to vote' against the household number or name, thus giving an indication of the level of registration in the stated constituency. However, the majority simply omit those not registered without providing any useful information or footnote.

Elector numbers and counts

Under the Representation of the People Regulations (England and Wales) Act 2001, numbering of the register should be sequential (in address order), although it recognises that additions and deletions could leave gaps. Each individual is given a unique number and all additions a suffix number to enhance identification. From our sample, the existence of non-sequential counts across the 100 wards was widespread. While the inclusion of an additional individual at the same address may reflect recent social changes, the increase in cohabitation or the higher turnover in student housing, there is little doubt that double entries (where two or more individuals or families are shown to be residing in the same place) exist, increasing the inaccuracy of the registers. As regards numbering, once again there seems to be a lack of uniformity. The majority of registers used the suffix number 1 and so on to indicate additions, e.g. unique numbers 1027/1 and 1027/2. By contrast, other wards used a variety of suffix numbers, which simply added to the confusion. For example, in one ward, four additions to a household of five people were given the unique numbers 1563/165, 1563/170, 1563/330 and 1563/660. Several wards simply failed to follow the regulations, inserting random numbers to indicate the inclusion of a new resident e.g. 37, 38, 2204, 39, etc. (Footnote – Regulation 41 doesn't permit this.) Furthermore, a few wards failed to adjust the ward electorate to account for the additional numbers. Street numbers, shown on the second page of the register, did

omit additions to the register, thus further compounding their inaccuracy. This is particularly important as they are often used to calculate the size of the electorate.

Recording of votes

Votes were recorded in a variety of ways. These included one or two line marks against the elector's unique number, ticks, crosses and lines through the whole name. Often a range of methods were used within the register and sometimes in the same polling district. Frequently this made it difficult to read and almost impossible to decipher who had voted or not. For instance, two lines through a name were also used to indicate that the person was deceased, while one line through a name meant that the elector had requested a postal vote.

A number of registers also included handwritten names of voters, unique numbers, street names and various comments ('moved', 'can't vote', 'no polling card'), which were intended presumably for the returning officer. Despite some registers including the comments 'can't vote' and 'no' against electors' names, a number were recorded as voting on polling day. Also, within a number of registers, whole streets had lines through them and there were references to the omission of voters because of error (inclusion of streets in the wrong polling district).

The final electronic version: step-by-step process

Initially, a stratified random sample of 100 wards was selected. For each selected ward, a hard copy of the marked roll was obtained and placed in an electronic version. Each marked roll contained a full name, address, postcode of each registered elector and a mark to indicate whether they voted in the 2001 Election or not.

The first part of the procedure involved coding each polling district within the ward and the ward itself. The polling districts were then placed into an overall register for the ward. At this stage there were some anomalies, which forced the removal of a number of wards. These included Rottingdean, Drynham (both stratum 1 – due to photocopying errors), Nash Mills (stratum 3 – incomplete register) and University (stratum 4 – incomplete register). Redcoat replaced University ward as it fulfilled the criteria for stratum 4 and was randomly selected using the same procedure as outlined earlier.

Other wards contained minor problems. Blythe Hill had 249 electors missing, while 91 electors in 'Town' ward in Halifax did not have any vote recorded due to a photocopying error. In both cases, these electors were removed from the numerator and denominator in calculations. In order to calculate the turnout rate, these electors need to be missing from the numerator and denominator so that we can use the turnout rate from the sample information available. Thus, when turnout rates are combined, the ward would be weighted by 249 electors less than there actually are.

Redwell ward suffered redistricting before 2001 and was now Redwell West and Redwell East. Following a discussion with Wellingborough Council and an examination of the Local Government Commission report, it was clear that no other changes were made. The two new wards made up the old Redwell ward. They both satisfied the sampling conditions and were therefore combined to form one ward as originally envisaged. In sum, the sample now contained 97 wards.

The main objective was to place the remaining 97 wards into one file in order to run the name recognition programs. Electors who are not eligible to vote in a general election were removed (marked by the letters G, F and L on the marked registers). This included electors who had their 18th birthday in the coming year but weren't eligible to vote on 7 June 2001. As stated earlier, some versions of the register contained first-time voters with electors' date of birth for the one-year cycle. In such cases, the letters FTV were included in the dataset. At this stage it was also necessary to delete name removed/ two lines crossed over elector/deleted elector from the register to account for redundancies caused by deaths.

Data from the 97 wards were inputted into an overall file. This overall file contained each elector's forename, surname, house number or address, postcode and town or city. Also included was whether the elector had voted or not, whether he/she was a proxy or postal voter, which ward and polling district the elector resided in, plus an identifier for which stratum the ward was sampled from and other geographical data such as the output area code. On some registers, the elector's date of eligibility was included where the elector was a first-time voter (we can tell this because the date was before 7 June 2001 and the elector is marked as voting). However, the majority of registers only record the date of eligibility to indicate that the elector could not vote in the election but can on the date stated or after the date. This had implications for defining first-time voters because many registers simply don't state whether they were first-time voters or not.

The total number of eligible electors was 585,617. However, 5,182 individuals who were registered but ineligible to vote at the 2001 UK Parliamentary Election were then added. A further 341 names were classified as missing data due to photocopying errors. Consequently, the total number of names in our overall register for the 97 wards was 591,140.

Notes

1 These are returned to the Department of Constitutional Affairs and retained for one year as a public record.

2 Unfortunately, 2001 Census data was not available at the time.

Appendix 2: Technical report – name recognition software

Using name recognition software to identify South Asians in the register

Nam Pehchan

Nam Pehchan is a computer program developed by Bradford Metropolitan Council and Bradford Health Authority for identifying names that originate from the Indian subcontinent and Sri Lanka (known as South Asian). The first version of Nam Pehchan (NP1) was developed in 1996 but was subsequently replaced in 2003 by an updated version, which is used in this project. Nam Pehchan 2 (NP2) was extended to cover non-Muslim populations more adequately, and to include Singhalese and Tamil names. A panel of language specialists has also checked all the names and their linguistic and religious affiliations. Names can now be assigned to one or two discrete languages and religions. This gives NP2 greater flexibility of interpretation and improved precision. Also, one of the main deficiencies of NP1 was that it focused on the Urdu-speaking Muslim population that predominated in Bradford and was weak in its interpretation of Hindi and Gujerati names. NP2 has amended its format to take account of this.

Nam Pehchan (NP2) identifies South Asian linguistic and religious origins of both surnames and forenames by matching against a stored list of names. The program attempts to match the full name or the name stem (the first five characters of an individual's name) so that it can provide a list of South Asians, including a language and religion origin for each person. The type of match ranges from 0–5 (no match to definitive match – see below) with a definitive match implying the most confidence in the match given:

0 no match on any name element

1 stem match on one name element, no match on others

2 full match on one name element, no match on others

3 full match on one name element, stem match on others

4 full match on all names

5 definitive match on at least one element.

There are seven religious codes (Hindu, Muslim, Buddhist, Sikh, other, common and clash), which the program assigns to a name along with the match. A 'clash' refers to 'inconsistent or contradictory data', while a 'common' code is assigned where the program identifies the name as South Asian but cannot determine the religious origin of the name. The Nam Pehchan software also ascribes each name with a gender code – one each for either male or female and the other where it cannot determine the gender from the name – and a language code of which there are 14. However, a number of the language categories were far too ambiguous, which prevented any meaningful analysis with the census categories and obstructed any comparison with previous academic outputs. We therefore decided to focus purely on religion.

Accuracy of Nam Pehchan

At the time of writing there hadn't been any rigorous assessment of Nam Pehchan version 2. However, its predecessor recorded levels in excess of 95 per cent for sensitivity, specificity and predictive value. Cummins *et al*. (1999) questioned these figures after testing out the accuracy of the Nam Pehchan version 1 program by using a dataset consisting of 356,555 cases of incident cancer between 1990 and 1992 in Thames, Trent, West Midlands and Yorkshire cancer registries. The results were compared with a reference standard, which included a combination of the program with visual inspection, itself facilitated through a computer-generated dictionary of South Asian names.

Cummins *et al*. (1999) found that Nam Pehchan version 1 recognised 5,506 cases as South Asian. The visual inspection highlighted 2,024 false positives (36.8 per cent) and 363 false negatives (9.5 per cent of those identified by the reference standard). There was also considerable geographical variation, with accuracy higher in cities such as Leeds and Bradford than those in London. For instance, in Yorkshire, NP1 had a sensitivity value of 96 per cent and a positive predictive value of 67.4 per cent. This compares with a sensitivity value of 88.2 per cent and a positive predictive value of 58.7 per cent in populations from the Thames region. Compared with the reference standard, NP1 had a sensitivity value of 90.5 per cent, but a positive predictive value of 63.2 per cent.

Cummins *et al*. (1999) concluded that Nam Pehchan version 1 identified a high proportion of names classified as South Asian by the reference standard, but there

was a high false positive rate. The authors recommend that it should not be used as a single strategy, but should be combined with dictionaries of common non-South Asian names.

Similarly, Mason *et al*. (2003) used NP1 to identify South Asian patients and brought in experienced experts in analysing South Asian names to manually check for any errors. They found that Nam Pehchan version 1 did not differentiate between different Muslim surnames (e.g. South Asian, Far Eastern or Arabic) and concluded that forenames (which were often more specific to language or region than surnames) needed to be examined to ascertain some South Asian people.

Characteristics of the sample using Nam Pehchan

Nam Pehchan version 2 identified 100,825 South Asians, which comprised just over 17 per cent of the sample population. More than half were identified as Muslims while around a quarter were either Hindu or Sikh. Within this, 778 names were identified as 'clashes', which were subsequently reassigned manually using an expert in South Asian names. However, for more than a quarter of the South Asian sample population, in excess of 25,000 names, the religious origin could not be determined. To combat this problem, we imputed actual religion for those classified by Nam Pehchan version 2 as 'religion not determined'.

However, during the reclassification of the 'religion not determined' category, it became apparent that a number of non-Asian names, particularly in predominantly non-Asian sampled wards, had been assigned to this category. Closer inspections of low-level matches also revealed a worrying number of false positives. In sampled wards with a large Hindu or Sikh population, there were markedly more South Asian names without an assigned religion. It seemed possible that Nam Pehchan was severely underestimating the number of Hindus in the sample population. To combat these problems, we decided to validate Nam Pehchan using alternative name recognition software and extensive manual checking.

SANGRA

In response to the inadequacies of Nam Pehchan version 1, Nanchchal *et al*. (2001) developed an alternative name recognition program. SANGRA (South Asian Names and Group Recognition Algorithm) incorporates directories of South Asian first names and surnames together with their religious and linguistic origin. The program was validated using health-related data with self-ascribed information on ethnicity.

Unlike Nam Pehchan version 2, the SANGRA code doesn't use stem names. The actual code used is provided below:

0 no match

1 full match on surname and forename

2 middle name match

3 forename match

4 surname match.

The academics who devised SANGRA claim that the match on the forename is more likely to indicate the family's ethnic identity with a particular linguistic or religious origin than the surname. However, during the manual check by an expert in South Asian names, it became apparent that the surname was a much better indicator of mixed Asians (examples of English forename and a South Asian surname throughout the dataset).

Regarding religious origin, SANGRA includes codes for Buddhist, Hindu, Muslim and Sikh. In cases where names could be either Hindu or Sikh, or even Christian or Hindu, a combination of these codes is assigned. This avoids any clashes or the reallocation of names into a 'religion not determined' category. The software also provides codes for ethnicity and language, although these weren't used during this study.

During a series of tests, Nanchchal *et al*. (2001) claimed that the SANGRA program had a sensitivity of between 89 and 96 per cent, specificity of 94 to 98 per cent, positive predictive value (PPV) of 80 to 89 per cent and negative predictive value (NPV) of 98 to 99 per cent. The figure for both Hindus and Muslims was 90 per cent or more for all categories, while only the sensitivity was below 90 per cent for Sikhs. There was also little geographical variation in the results across the UK. In sum, Nanchchal *et al*. (2001) concluded that SANGRA is:

> … a valid method of ascertaining South Asian origin by name and, to a lesser degree of accuracy, of differentiating between the main religious and linguistic subgroups living in Britain.

Characteristics of the sample using SANGRA

SANGRA identified 90,447 South Asians, just over 15 per cent of the sample population. This was more than 10,000 names less than those classified by Nam Pehchan version 2. Once again, more than half were identified as Muslim, although the total number of Muslims in the sample was around 7,000 less than those identified by Nam Pehchan. When SANGRA was used, the number of Hindus in the sample more than doubled, adding credence to the suspicion that they had been misclassified by Nam Pehchan. The number of Sikhs was slightly lower, although this partly reflected the 4,229 names that were jointly classified Hindu and Sikh. In total, SANGRA jointly classified 7,815 names.

Adjusting the NP2 sample using SANGRA and manual checking

For the calculation of registration and turnout it was necessary to obtain the most accurate recognition of South Asian names. Table A2.1 shows a cross-tabulation of the total South Asians identified by Nam Pehchan and SANGRA by their respective match codes. Given this information it was possible to decide which particular cells would be manually checked and where SANGRA could be used to validate Nam Pehchan. We decided to use Nam Pehchan as the main source given that it was used to provide initial turnout findings. Also evidence from Steele (2005), who obtained a dataset from Coventry Primary Care Trust to compare the two programs against a list of names, suggests that Nam Pehchan version 1 had a higher predictive value than SANGRA. Given that we were using a more updated version of the Nam Pehchan software, it was decided that SANGRA should be mainly used as a validation tool.

Table A2.1 Classifications of names from our sample by Nam Pehchan and Sangra according to the level of match

NP/SANGRA match	0	1	2	3	4
0	485,855	247	0	581	2,739
1	2,407	92	0	158	610
2	91	10	0	10	43
3	8,395	3,600	0	3,508	13,184
4	2,354	11,783	0	1,471	15,087
5	1,250	16,269	0	1,968	19,087

No cases in SANGRA code 2 because few middle names on the register were entered onto the data file.

Table A2.2 shows the matrix used to determine the South Asian sample. First, where Nam Pehchan identified a full (match 4) or definitive (match 5) match and SANGRA identified the name as South Asian, Nam Pehchan was used to assign the religious origin of the name. Through our own and expert analysis of the names, we were confident that Nam Pehchan had correctly classified full and definitive matches. In total, nearly two-thirds of the original South Asian sample (65,665 names) from these cells was assigned their original classification from Nam Pehchan. Also, where SANGRA assigned a full match on both the forename and surname (match 1) and Nam Pehchan provided a full match or a stem match on one name element (match 1, 2 and 3), we decided to use the SANGRA code. These 3,702 names were added to a non-manual file for South Asians, totalling 69,367 names. It included cells only where we had already decided what software to use. In circumstances where the religion could not be determined, the other software was used to verify the religion. Where both Nam Pehchan and SANGRA couldn't determine the religious origin of the name, we adjusted manually according to the judgements of the South Asian expert.

Table A2.2 Matrix by which names from our South Asian sample were assigned

NP/SANGRA match	0	1	2	3	4
0	X	MAN	0	MAN	MAN
1	X	SANG	0	MAN	MAN
2	X	SANG	0	MAN	MAN
3	MAN	SANG	0	MAN	MAN
4	MAN	NP	0	NP	NP
5	MAN	NP	0	NP	NP

No cases in SANGRA code 2.

Second, where neither Nam Pehchan nor SANGRA provided a full match, with the exception of 247 names that Nam Pehchan classified as non-Asian and SANGRA South Asian, names were manually checked by a South Asian name expert (see Table A2.2). This manual file contained 33,079 names. However, before the manual check took place, we coded identifiable English names to non-Asian, which reduced the number of names checked to around 30,000. There were a number of problems encountered during the manual checking process. For instance, it was noticeable that a large number of South Indian or Sri Lankan names were concentrated in particular wards. While these were mainly coded as Hindu, it was decided that some would be assigned to the 'other South Asian' category given that they were likely to be Sri Lankan. We also identified some Parsis in the dataset. These were reclassified as Hindus. The growing number of mixed marriages also posed some problems. Occasionally the surname was English/Indian (Hindu/Muslim/Sikh), which made it difficult to assign the religion, particularly when the forename was English. It

was also clear that a growing number of younger Hindus were adopting Muslim names. This not only caused problems when manually determining the religious origin of the name but undoubtedly contributed to the large number of names that were either coded by the name recognition programs as 'religion not determined' or were classified as having more than one religion.

Of the 590,799 names in the sample population (excluding the 341 missing cases), 488,353 names were not examined. This included the 485,855 names that were identified as non-Asian by both Nam Pehchan and SANGRA. Previous inspections of the data revealed that the remaining 2,498 names that were assigned a low match (stem match or full match on one name element) by Nam Pehchan were predominantly non-Asian. These were reclassified accordingly.

Religious origin of sample population

Table A2.3 shows the religious origin of the sample population following the validation process. Around 100,000 names were identified as South Asian, with Muslims representing more than half the South Asian sample. However, there were significant increases in the number of Hindus and Sikhs, confirming previous suspicions that the vast majority of names where the religion was not determined fell into either of these two categories. Names of Sri Lankan origin largely made up the 'other South Asian' category. We used this validated classification to calculate turnout and registration across and within different South Asian communities.

Table A2.3 Religious origin of sample following validation

Religion	Validated unweighted sample
Hindu	26,891
Muslim	54,452
Sikh	16,901
Other South Asian	1,045
All South Asian	99,289
Non-Asians	491,510
Total	590,799

Note: the total number excludes 341 missing cases. The total number includes 5,182 names not registered to vote in UK parliamentary elections.

Appendix 3: Technical report – method used to estimate registration

Registration

Issues with the 2001 Census

It is not possible to estimate registration for the complete set of 97 wards at the ward level because of boundary changes. The Census Area Statistics (CAS) wards are based on 2003 boundaries; so any redistricting that occurred after 1 May 2003 is not relevant for this study. Our study uses 2001 wards, yet there was significant redistricting in 64 of the 97 wards between June 2001 and 1 May 2003. Some wards also underwent changes before 2001, which may not have been taken into account by our sample – for example, Freshbrook ward in Swindon South was in fact Freshbrook and Grange Park in 2001.

Estimates of the number of registered electors and information on whether they voted are derived from our sample of marked electoral registers (see above). The census output area (OA) is the smallest level of analysis for which we can obtain estimates of population, registered electors and whether these electors voted. This is therefore the most appropriate level of analyses to examine geographical variations. However, in order to achieve global estimates of registration by religion, it is simple to aggregate OA areas across the entire sample. In other words, OAs are the building blocks to generate a VAP/VEP to compare with our sample of registered electors, rather than a unit of analysis in their own right.

The number of registered electors of South Asian and other origins was then compared with the relevant census population. This was reasonable due to the date of the finalisation of the electoral registers for the 2001 General Election and the census enumeration date (29 April) being within a few weeks of each other. Adjustments were made to take into account the number of attainers reaching 18 by election day and also the projected number of deaths by area. When making estimates of registration it was also necessary to make adjustments for redundancies in the register and for the estimated number of adults born outside of countries eligible to vote in any UK elections (i.e. outside the UK, Europe and the Commonwealth). This is detailed further below.

Constituency estimates of registration were also made based on constituency electorates (2001) and the census populations. The aim was to combine estimates of turnout with estimates of registration to make a more realistic estimation of participation than analysis of turnout alone.

Matching postcodes to census output areas

There were 3,192 output areas for the 97 wards, each containing separate totals for each religious category (non and other Asian, Muslim and non-Muslim South Asian electorate per OA) and the overall electorate. Levels of electoral registration in 2001 were estimated for valid OAs in our sample, derived from postcodes using the AFPD, which matches postcodes to other geographies. Not all postcodes in our sample were included since some fell into output areas with incomplete coverage (see above). There were also a number of unmatched overseas electors who appeared on the register but did not have a postcode. These were excluded from our analyses, as, aside from having no postcode, they should not appear in the denominator that covers only UK residents (see below).

Initially, 570,014 (97.3 per cent) of the 585,617 cases were matched, leaving 15,603 (2.7 per cent) unmatched. Some of the problems with the unmatched postcodes were solved relatively simply. For instance, some postcodes contained only five characters and a review of the postcode software and the hard copies of the registers revealed the correct postcodes. Some postcodes were incorrect and had to be amended using the methods described above. However, the remaining unmatched cases posed more serious problems. First, parts of some wards simply did not have postcodes – examples include Ilfracombe Central and Swimbridge – both on the marked registers and on the UK postcode data file. In such circumstances, we had to contact the electoral services in these areas to obtain local copies of the postcode.

Second, a number of unmatched postcodes stem from overseas electors who were included on the electronic version (from the marked register) but didn't have any postcode. We simply regarded these as missing.

Those 5,182 electors who were ineligible to vote at the 2001 General Election but were on the electoral register were then added into the overall file and the postcodes were matched to the appropriate census output area.

Calculating registration

Creating the numerator

To calculate registration we used the validated religious origin of names from our sample as the numerator. To allow comparison with census categories (the denominator), the religion variable was subdivided into the following categories (for full definitions see Table A3.1 later in this appendix):

1 Muslim (all Muslims, plus other Pakistanis and Bangladeshis with no recorded religion)

2 non-Muslim South Asian (all Asian or mixed white and Asian Hindus and Sikhs, plus Indians, Pakistanis and Bangladeshis who are not Muslim)

3 non-South Asian (all other categories).

Key variables including the validated religion variable recoded into the above categories were then aggregated from our individual file to the output area level. There were 3,192 output areas for the 97 wards, each containing separate totals for each religious category (non and other Asian, Muslim and other South Asian electorate per OA) and the overall electorate. This file was then merged with our postcode file, which included the number of postcodes per OA from the electoral register and alternatively from the 2001 Census figures. As noted above, we calculated the percentage postcode coverage for each OA and, of the 3,192 OAs, 1,823 had 100 per cent postcode coverage. To be certain of accurate registration rates, we include OAs in our analysis only where the postcode coverage of our sample had 100 per cent coverage and matched that of the Census as indicated in the AFPD by the count of valid residential postcodes falling in each 2001 Census OA.

Creating the denominator

The first step in the creation of the denominator was to estimate the total population aged 18 (VAP) by OA, making adjustments for the estimated numbers of attainers and deaths (calculated using census information based on the 39 days from the census date to the general election date). Similar estimates were generated by ethnicity and by religion using the ethnicity and religion census tables at the OA level. We then adjusted these OA totals to take into account ineligible electors, using information about country of birth (together with ethnicity and religion) at the ward level (see Box A3.1).

Box A3.1 Eligible countries used in the COB adjustment

- United Kingdom
- Republic of Ireland
- Channel Islands
- All European Union countries
- Nigeria
- Countries in South and East Africa
- Countries in South Asia
- Canada
- Caribbean and West Indies
- Oceania countries

Estimated ward-level rates of ineligibility were applied evenly across all OAs within each ward. Using the resultant OA religion and ethnicity (marginal) totals together with ward-level ethnicity by religion cross-tabulations, we created estimates for each ethnic-religious subgroup at the OA level. These estimates were constrained to match the ethnicity and religion marginal totals. These were then summed to the major subgroup categories – non and other Asian, Muslim and non-Muslim South Asian – and constrained to the OA VAP. Table A3.1 shows which constrained ethnic and religious group from census standard table ST104 ('Ethnicity by religion') was assigned to each subgroup category.

The final step was to calculate the total electorate (total voting-age population) on polling day after taking account of attainers (calculation using census information based on the 39 days from the census date to the general election date) and deaths. It must be noted that adjustments for attainers and deaths are made only at the total level, unlike ineligibles (e.g. COB) for which the adjustments are religion specific. Yet, for us to create denominators with country of birth adjustments (total voting eligible population), we multiplied the total voting-age population by the proportion of all people from eligible countries. We then summed all our registration categories (non and other Asian, Muslim and other South Asian) to the total voting eligible population on polling day. This provided us with three voting eligible denominators for each of our registration categories, plus the total voting eligible population on polling day. The latter could be used as the denominator in the calculation of overall registration. We did encounter problems with negative numbers among the counts –

we adjusted these numbers to 0 – and the rounding issue concerning 0s and 3s at the output-area level. However, these were dealt with early on in the creation of the denominator.

Two files were created containing numerator and denominator categories, one with and one without country of birth adjustments. Both files included only OAs with 100 per cent postcode coverage. To obtain registration rates, we simply divided the sum of all the numerators (across all OAs) by the sum of the denominators across the same areas, weighting where necessary. As discussed in Chapter 2, a number of different issues were considered, including and excluding areas with denominators less than 6.5. Because of the unreliability of estimates in areas with small denominators (and their large weights due to the sample design) we decided to restrict our analyses to areas with denominators greater than 6.5.

Table A3.1 Ethnicity by religion, 2001 Census table ST104: matrix of subgroups assigned to registration categories

Ethnicity/ religion	**Christian**	**Buddhist**	**Hindu**	**Jewish**	**Muslim**	**Sikh**	**AOR**	**No religion**	**RNS**
White	NA	NA	NA	NA	M	NA	NA	NA	NA
WBC	NA	NA	NA	NA	M	NA	NA	NA	NA
WBA	NA	NA	NA	NA	M	NA	NA	NA	NA
WA	NA	NA	OSA	NA	M	OSA	NA	NA	NA
OM	NA	NA	OSA	NA	M	OSA	NA	NA	NA
Indian	OSA	OSA	OSA	OSA	M	OSA	OSA	OSA	OSA
Pakistani	OSA	OSA	OSA	OSA	M	OSA	OSA	M	M
Bangladeshi	OSA	OSA	OSA	OSA	M	OSA	OSA	M	M
Other Asian	NA	NA	OSA	NA	M	OSA	NA	NA	NA
All black	NA	NA	NA	NA	M	NA	NA	NA	NA
Chinese and OEG	NA	NA	NA	NA	M	NA	NA	NA	NA

Key: subgroup categories
NOA = non and other Asian subgroup category
M = Muslim subgroup category
OSA = other South Asian subgroup category
Other details
AOR = any other religion
RNS = religion not stated
WBC = white and black Caribbean
WBA = white and black African
WA = white and Asian
OM = other mixed
OEG = other ethnic group

Appendix 4: Technical report – method used to estimate turnout

Calculating turnout: adjustments to the register

Before estimating turnout, we made a number of adjustments to the marked registers. First, we excluded all those on the register who were not eligible to vote in parliamentary elections. These included foreign nationals eligible to vote in local elections and young people who reached the age of 18 during the life of the register but were still 17 on 7 June 2001 (attainers). Electors who had applied for a postal vote (3.9 per cent in our sample) are excluded, as we have no evidence as to whether these votes were cast. Only 10.6 per cent of postal voters in our sample were South Asian. This represented 2.4 per cent of our overall sample of South Asians. The number of postal voters in our sample is similar to the national figure in 2001. In this paper we focus on the relative propensity to vote among non-postal voters. Following a change in the electoral rules, postal voting increased substantially at the 2005 General Election to around 15 per cent. Proxy voters are included, as their votes are marked off at the polling station in the normal way.

Weighting

Because of the complex nature of the sample design, estimates of sampling variance should take into account the use of wards as primary sampling units, the stratification of wards and the use of weighting, including post-stratification weighting.

The sample contains wards from England and Wales, and excludes postal voters (these are regarded as missing because, at the individual level, we do not know if the votes were cast). Postal voters are also removed from the population for England and Wales. This is achieved in the following way.

A = total electorate (39,227,923) – number of postal votes issued (1,758,000)

A= 37,469,923

B = total number of votes cast (23,243,308) – postal votes cast (1,370,884)

B = 21,872,424

C = A (37,469,923) – B (21,872,424) (non-votes)

C = 15,597,499

Target population: 21,872,424/37,469,923 (B/A) = 58.4 per cent.

Weights are applied to reflect the stratification of the sample of wards (the design weight 'Gweight'). We constructed the Gweight as follows:

Gweight = number of wards in strata population/number of wards in strata sample.

Stratum 1 = Gweight = (7191/18).

Stratum 2 = Gweight = (1972/20).

Stratum 3 = Gweight = (1025/19).

Stratum 4 = Gweight = (201/20).

Stratum 5 = Gweight = (163/20).

The 2001 turnout in England and Wales excluding postal voters is 58.4 per cent. However, initially, stratum 0 (2,057 wards with no South Asians) is included in the population but not in the sample. We therefore recalculated the Gweight on the assumption that the population, which by definition is non-Asian, behaves as the non-Asian population in stratum 1. In effect, the 2,057 wards in stratum 0 were added to the 5,134 wards (0<0.5 per cent South Asians in ward) in stratum 1.

Although we have a sample of over half a million voters, the sample is clustered in 97 wards and this will be reflected in the sampling errors. It is not surprising, therefore, that the total turnout estimated from our sample, at 56.6 per cent, differs from the actual turnout rate for England and Wales in 2001 of 58.4 per cent. This reflects the sampling variation due to the selection of wards.

We can adjust for this by applying a weighting factor so that the total number of voters and non-voters estimated from the sample equals the actual numbers of voters and non-voters in England and Wales (the design and vote weight 'Vgweight'). The key stages in the calculation of the correction weight are shown below.

X = sample registered – exclude postal voters.

Y = sample vote – exclude postal voters.

These figures are grossed up from weighted by strata sample ward *n*/strata population ward *n* (using Gweight). Vgweight is therefore grossed up to the actual population and adjusted for actual turnout.

Non-voters: 23,677,659
Voters: 30,832,946
Electorate: 54,510,605 (excludes postal voters)

For Vgweight = B/Y (votes) C/X (non-voters)

The non-vote weight:
15,597,499 (England and Wales)/23,677,659 (our sample) = 0.6587 (correction factor).

The voted weight:
21,872,424 (England and Wales)/30,832,946 (our sample) = 0.7094 (correction factor).

However, use of the correction factor assumes that the stratified sample of 97 wards was no more or less representative of the South Asian population (which it was designed to represent) than the overall population. It may or may not represent South Asians more accurately than the rest of the population. The key point is that we assume that the lower turnout rate in the sample has affected Asians and non-Asians equally. For completeness we present estimates with and without this additional weighting factor. The results are reported in Chapter 4.

Appendix 5: Technical report – variables used in models of registration

Output area

OA per cent overall agriculture OA level
OA per cent overall manufacturing OA level
OA per cent Hindu religion (main effect)
OA per cent Muslim religion (main effect)
OA per cent Sikh religion (main effect)

OA long-term ill (religion specific)
OA owner-occupation (religion specific)
OA pensioners (religion specific)
OA unemployment (religion specific)

Ward

Ward ID
Ward per cent overall lived at the same address (inverse of migration within UK)

Ward owner-occupation (religion specific)
Ward degree (religion specific)
Ward full-time students (religion specific)
Ward NSEC 1 and 2 (religion specific)
Ward NSEC 6 and 7 (religion specific)
Ward two cars (religion specific)
Ward Indian (religion specific)
Ward long-term ill (religion specific)

Constituency

2001 General Election: targeted by any of the three parties
Margin 1997

Appendix 6: Technical report – variables used in multilevel modelling

Individual level

Vote
Religion
Person ID – individual level

Household level

Household ID
Number of people in household

Output area

OA ID
OA per cent overall agriculture OA level
OA per cent overall manufacturing OA level
OA per cent Hindu religion (main effect)
OA per cent Muslim religion (main effect)
OA per cent Sikh religion (main effect)
OA long-term ill (religion specific)
OA owner-occupation (religion specific)
OA pensioners (religion specific)
OA unemployment (religion specific)

Ward

Ward ID
WARD per cent overall lived at the same address (inverse of migration within UK)
Ward owner-occupation (religion specific)
Ward degree (religion specific)
Ward full-time students (religion specific)

Ward National Statistics Socio-economic Classification (NS-SEC) 1 and 2 (religion specific)
Ward NS-SEC 6 and 7 (religion specific)
Ward two cars (religion specific)
Ward Indian (religion specific)
Ward long-term ill (religion specific)

Constituency

2001 constituency reference
2001 General Election: targeted by any of the three parties
2001 party spending (three parties)
Total size of electorate in 2001
Percentage majority 1997
Ethnic candidate
Gender of candidate
Incumbent candidate or new candidate
Votes cast as a proportion of the electorate, 1997
Margin 1997
Constituency ID alphabetical order