Cathie Marsh Centre for Census and Survey Research

Registration of South Asian Populations at the 2001 General Election: New Evidence from the 2001 Census.

CCSR Working Paper 2006-07
Ed Fieldhouse, David Cutts and Paul Norman.
Ed.fieldhouse@manchester.ac.uk, David.Cutts@manchester.ac.uk,
Paul.Norman@manchester.ac.uk

This paper reports research based on analysis of the 2001 Census and a sample of marked electoral registers. These data allow us 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. We make three important findings. First, unlike turnout, South Asian registration is lower than that of the population as a whole. Second, a large proportion of the difference in registration rates is attributable to ineligibility associated with country of birth. Third, South Asian registration is higher in areas where South Asians are most geographically concentrated.

www.ccsr.ac.uk

Registration of South Asian Populations at the 2001 General Election: New Evidence from the 2001 Census.

Edward Fieldhouse, David Cutts and Paul Norman

Cathie Marsh Centre for Census and Survey Research, University of Manchester

Abstract

This paper reports research based on analysis of the 2001 Census and a sample of marked electoral registers. These data allow us 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. We make three important findings. First, unlike turnout, South Asian registration is lower than that of the population as a whole. Second, a large proportion of the difference in registration rates is attributable to ineligibility associated with country of birth. Third, South Asian registration is higher in areas where South Asians are most geographically concentrated.

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 aged population or VAP (Lijphart, 1997), although since 2001 the voting eligible population or VEP estimate is now used by a number of leading US scholars (see McDonald & Popkin, 2001). The VEP estimate takes account of adjustments for those born outside of eligible countries as opposed to the non-adjusted VAP. In most European countries the denominator for turnout calculations is the registered electorate, which can be as much as 7% 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 under coverage, 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 responses against the electoral register each and by providing booster samples of minority electors (e.g. the 1997 British

Election Study (BES) 'Black and Minority Ethnic' booster sample). However, despite the undoubted value of such surveys, the sample size problem remains. For instance, the 1997 BES 'Black and Minority Ethnic' booster sample only contains 227 Asians of Indian origin and 124 Asians of Pakistani origin. Alternative approaches include ecological or area based estimates but these are 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 paper we use information from a complete set 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 British South Asian communities. In 2001 the General Election (June 7th) and Census day (April 29th) were remarkably close. The close co-incidence of an election and a census provides a unique opportunity to undertake analysis of registration at a time when voter apathy was a key election issue. Registration is assessed here by comparing the census population with our sample of marked electoral registers from the 2001 General Election. These are analysed using names recognition software which is able to identify names with a South Asian origin (i.e. from the Indian sub-continent). Together with geographical population information from the 2001 UK Census, this information allows a unique analysis of registration amongst Britain's South Asian communities in 2001. We pay particular attention to ensuring that the denominator and numerator used in rates closely correspond by population subgroup with detailed allowances made for demographic change between the time points noted above.

UK Registration

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. Like most other western democracies, the UK voter qualification age is 18. In the UK, registers are compiled by local authorities who 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 those 17 year old 'attainers' 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. At the 2001 General Election rolling registration was introduced. 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 June 7th, new electors were required to register before 5th April. The system also allows individuals to update their details during a particular year. This led to a 1.3% increase in the number eligible to vote in 2001 compared to 1997 (Electoral Commission, 2001).

In the UK, evidence from comparing the 1991 Census and the Post Enumeration Survey estimated that 7.1% of the people eligible to vote were not on the electoral register (Smith, 1993). A later study estimated that 4.8% of people enumerated in the 1991 Census were not on the electoral roll (Heady et al, 1997). 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 3000 (Johnston et al, 1997). At the 2001 general election, one study estimated registration at just under 97% (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% in England and 99.4% in Scotland (Dorling, 2007). When compared to 2001 census figures, it was estimated that around 7% of people in England and Wales were not on the electoral register in 2002. However, the figures should be treated with caution given the uncertainty of population estimates in London and the North West (Dorling, 2007). An estimate for 2004 suggests that UK registration rates might range from 92-93% 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 still relatively small.

Variations in Registration

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% higher for men than women and found levels of non-registration to be higher for the youngest age-groups (17 year old 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 ethnic

minority groups (Anwar, 1994; 1998; Smith 1993; Saggar, 1998a). 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% for Indians, 90.2% for Pakistanis, 91.3% for Bangladeshis and 96.9% for Whites (Saggar, 1998a). One possible explanation of this is that levels of British citizenship vary across BME groups with the highest levels of non-citizenship among Black Africans and Bangladeshis. More recently, a face to face sample survey across five local authority areas found non-registration levels to be higher among Indians (24%) than Whites (18%), Pakistanis (17%) and Bangladeshis (13%) (Anwar, 1998a). The Electoral Commission/ONS study discussed above suggested that the percent not registered in South Asian communities was much lower: 6% for Indians and Bangladeshis, and 8% for Pakistanis, compared to 17% for all British Minority Ethnic groups and 6% for whites. The study also found that there was a strong relationship between non-registration and nationality, which, as will be seen below, is also reflected in our results.[1]

Various explanations have been put forward as to why registration rates vary for minority groups. In part, this may be dependent on the methods used by electoral registration officers and diverse local authority policies on updating the register (Smith, 1993). It is also claimed that registration offices have not sufficiently changed their practice to meet the demands of the BME electorate (Anwar, 1990; 1998). Certain factors such as language difficulties, unease about dealing with officialdom, concerns with anonymity, fear of harassment and doubts about residence status disproportionately affect BME communities more than the wider population and have

---

[1]Due to the possibility of dual and acquired citizenship, ONS assumed that people born in countries outside of the Europe and the commonwealth were eligible to be registered.

contributed to varying levels of non-registration (Anwar, 1990; 1996, 1998a). Indeed, survey evidence from Bradford found that deliberate non-registration among Asians was much lower than other BME groups (Le Lohe, 1990). Whilst we do not set out to explain registration, but rather to measure it, we do suggest that nationality does play a significant part, as too does the geographical distribution of South Asian groups. In measuring these effects we also provide some insight into other factors associated with registration.

Calculating registration rates at the constituency level

After obtaining the registered electorate for each constituency for England and Wales, we derived the Voting Age Population (VAP) from 2001 Census data. 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 2001 general election (VAP) whereas only 39,205,725 people were registered to vote. The estimated registration rate for England and Wales was 97.25%.

Not surprisingly, there were wide spatial variations in estimated registration rates. A number of constituencies recorded registration rates above 100%; in other words more adults were registered to vote than there were adults in the population to register! This discrepancy could be explained by a census under-enumeration (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.

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 (-.267) between South Asians and registration at the constituency level suggesting low turnout for South Asian electors. But does this ecological relationship hold at the individual level? Whilst 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 % South Asian (2001 Census Data); Correlation Coefficient (-.267)
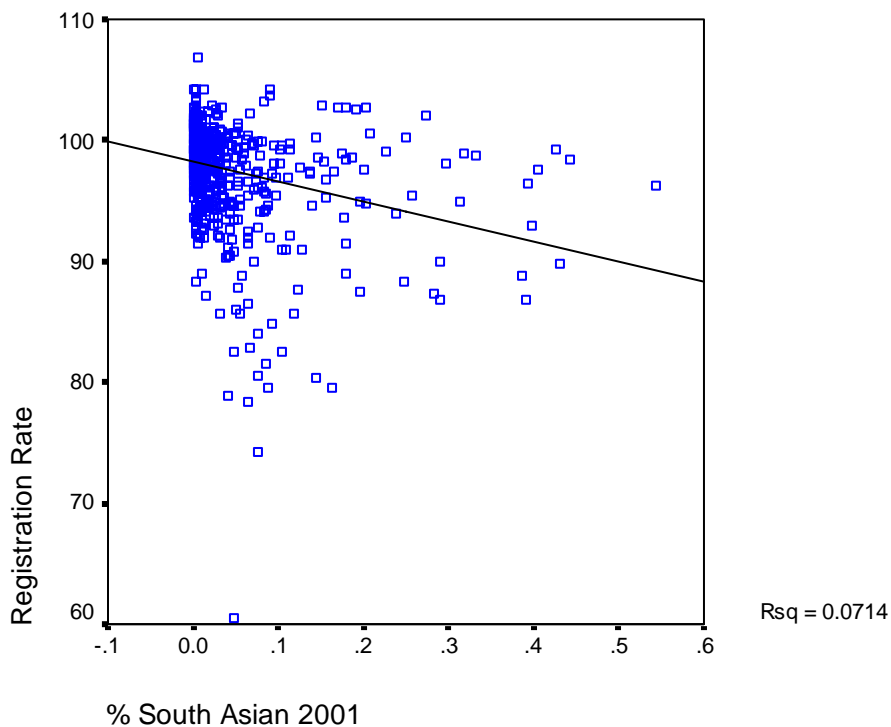
Generally the constituency registration rates provide a 'ball park' guide to the level of registration across England and Wales. However because they are based on aggregated official electorates no adjustments can be made for the existence of ineligible electors on the registers. Not only do some of 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.

Method

At all general elections, the electoral registers are manually marked according to whether each registered voter actually voted. As noted above, this research uses marked registers from the 2001 general election, for a sample of ninety seven wards, based on a stratified random sample. Wards were sampled disproportionately in areas with a large South Asian population to ensure the effective coverage of different subgroups. Stratification weights are applied to make the results nationally representative. All electors were included in the selected wards, which were used as the primary sampling units (see table 1). The sample contains wards from England and Wales and includes postal voters.

Table 1. **Stratified Random Sample**

	% South Asian Pop.	No of Wards	Sample	No of Sampled Electors	No of Sampled S.Asian Electors
Group A	0% - <0.5%	7191	18	98856	1009
Group B	0.5% - <2%	1972	20	87996	2079
Group C	2% - <10%	1025	19	114266	7585
Group D	10% - <20%	201	20	130164	25112
Group E	20% +	163	20	154335	65040
Total		10552	97	585617	100825

Note. All registered electors in the sampled wards are included,. The original sample was 100 wards, but due to data problems three wards were excluded.

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). Our study uses 2001 wards, yet there was significant redistricting in 64 of the 97 wards between June 2001 and 1st May 2003. Consequently we do not attempt to calculate registration at the ward level. Except for those 33 unchanged wards, the 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 analysis to estimate registration and examine geographical variations. As a unit of analysis they have 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, OA's also provide the building blocks to generate an aggregate VAP/VEP to compare with our sample of registered electors.

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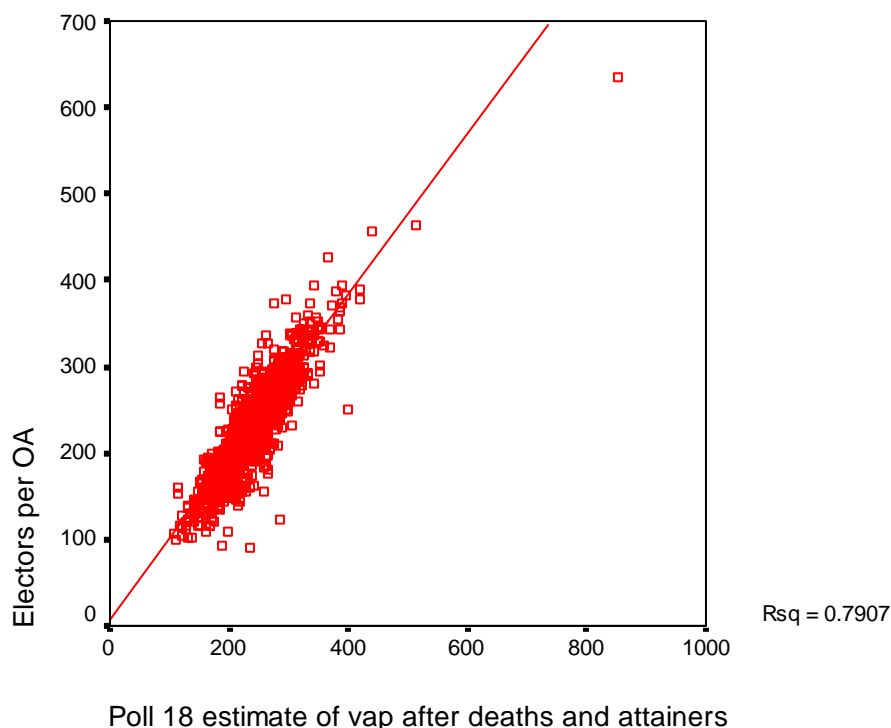

Poll 18 estimate of vap after deaths and attainers

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 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.[2] 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. However some of the differences

cancel each other out and 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

---

[2] 97% of electors were successfully allocated a postcode, and 1823 out of 3192 OAs were retained as valid under the criteria described. This is described in more detail below.

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.

Identification of South Asian Electors

Religious origin was derived from electors' names using 'Nam Pehchan' (NP) and SANGRA. NP uses a validated dictionary of names common in Britain, but originating from the Indian subcontinent and Sri Lanka (known as South Asian). The software identifies South Asian linguistic and religious origins of both surnames and forenames by matching against a stored list of names. The programme 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. However, a number of the language categories were far too ambiguous which prevented any meaningful analysis with the 2001 Census categories and obstructed any comparison with previous academic outputs. We therefore focus on religion.

An updated version of Nam Pehchan (NP2) is used in this paper. At the time of writing there hadn't been any rigorous assessment of Nam Pehchan version 2 (NP2). However, its predecessor recorded levels in excess of 95% for sensitivity, specificity and predictive value (Harding et al, 1999). Others have questioned this, suggesting more substantial rates of misclassification (Cummins et al, 1999). To combat this problem, NP2 was extended to cover non-Muslim populations more adequately and to include Singhalese and Tamil names, as well as improving the identification of Hindi and Gujerati names. Names can now be assigned to one or two discrete languages and

religions which gives NP2 greater flexibility of interpretation and improved precision. It is widely considered to be the most reliable method for identifying Asian names.

Despite this, we rigorously checked the data for possible misclassifications and noted that a number of Non-Asian names, particularly in predominantly Non-Asian sampled wards, had been assigned a religion not determined category. Closer inspections of low level matches also revealed a small number of false positives. In sampled wards with a large Hindu or Sikh population, there were noticeably more South Asian names without an assigned religion. It seemed likely that NP2 was slightly underestimating the number of Hindus in the sample population. To combat these problems, we validated the classification using an alternative name recognition software SANGRA (South Asian Names and Group Recognition Algorithm) and extensive manual checking.

Nanchchal et al (2001) developed SANGRA in response to the inadequacies of NP1. It 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. For both Hindus and Muslims, sensitivity specificity and predictive value were in excess of 90%. The sensitivity was below 90% only for Sikhs. There was also little geographical variation in the results across the UK. Given evidence from Honer (2004) who compared the two programs and found NP1 to have a higher predictive value than SANGRA, it was decided that SANGRA should be mainly used as a validation tool.

In cases where NP2 and SANGRA did not agree on the religious origin of the name religion was assigned manually, by an expert in South Asian names. Despite these rigorous efforts, there still remains the possibility of misclassification in ascribing religion in our analyses. Such misclassifications would be akin to measurement error in a survey.

Table 2. Religious Origin of Sample following Validation

Religion	Validated Unweighted Sample
Hindu	26891
Muslim	54452
Sikh	16901
Other South Asian	1045
All South Asian	99289
Non Asians	491510
Total	590799

The total number excludes 341 missing cases and includes 5182 names not registered to vote in UK Parliamentary elections

Table 2 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. Names of Sri Lankan origin largely made up the Other South Asian category.

Constructing the Numerator

The wide range of differences and inconsistencies between electoral registers meant that significant adjustments were required before any analyses could take place. For instance, some registers still included 'deleted electors' (removed from the register because they may have moved house or died) in the ward electorates. A number of registers also included additional names entered by hand which were not included in

the ward electorate. From our sample, the existence of non-sequential counts across the 97 wards was widespread. Whilst 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 was evidence of double entries (where two or more individuals or families are shown to be residing in the same place) increasing the inaccuracy of the registers. Also while suffix numbers were used to indicate additions, a few wards failed to adjust the ward electorate accordingly. In every case we made adjustments to the registers to take account of wastage and additions so that we could obtain more reliable ward electorates.

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 sub-divided into the following categories:

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)

There were 3192 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 match postcodes to other geographies. Not all postcodes in our sample were

included since some fell into output areas with incomplete coverage (see above). There was 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 which covers only U.K. residents (see below).

As noted above, we calculated the percentage postcode coverage for each OA and of the 3192 OA's, 1823 had 100% postcode coverage. To be certain of accurate registration rates, we only include OAs in our analysis where the postcode coverage of our sample had 100% 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.

Constructing 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.[3] 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 sub-group at the OA level. These estimates were constrained to match

---

[3] Eligible countries used in the country of birth adjustment were as follows: United Kingdom, Republic of Ireland, Channel Islands, all European Union countries, Nigeria, countries in South and East Africa, countries in South Asia, Canada, Caribbean & West Indies and Oceania countries.

the ethnicity and religion marginal totals. These were then summed to the major sub-group categories – Non and Other Asian, Muslim and non-Muslim South Asian, and constrained to the OA VAP.

2001 South Asian Registration Rates

Table 3 shows the estimated registration rates for all our sampled OAs (unweighted) by the identifiable religion groups, before and after adjustments for country of birth. 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).[4]

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. Whilst 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

---

[4]Empirical analyses show that cells with counts of six 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 mid way between 6 and 7.

the Muslim population. This inevitably affects registration rates. For example, the Electoral Commission research showed that by non-registration amongst Muslims living in the U.K for ten years or more was only 6% compared to 14% amongst all Muslims.

Table 3.   Registration by religion, (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	With COB adjustment (Unweighted)	Before COB adjustment (Unweighted)	With COB adjustment (Weighted)	Before COB adjustment (Weighted)
Overall	100.6% (1823)	96.8% (1823)	100.4% (1823)	98.5% (1823)
Non & Other Asian	101.5% (1823)	98.1% (1823)	100.5% (1823)	98.9% (1823)
Non-Muslim South Asian	92.6% (754)	90.6% (763)	90.7% (754)	88.1% (763)
Muslim	96.4% (896)	89.5% (944)	91.5% (896)	82.0% (944)
All South Asian	95.5% (1150)	90.7% (1182)	93.4% (1150)	86.2% (1182)

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

Table 3 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 is 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%, higher than the non-Muslim South Asian rate of approximately 93%. The non-Asian unweighted rate after country

of birth is adjusted exceeds 100% 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). Whilst 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 predominantly Hindu and Sikh other South Asian group, 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 sub-groups (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 3.

The effect of the proportion of people within each religion 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 (see table 4 below). The dependent variables are the OA registration rates for each religion category, and the explanatory

variables are the percent within each religion group who were born in ineligible countries (i.e. the 'religion specific' rate of ineligibility). Table 4 shows there is a decrease in registration across all groups as the percentage born outside eligible countries increase, and that the rate of decrease is smaller for Non-South-Asians. Whilst 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.

Table 4. OLS Regression coefficients of OA registration – Country of Birth

Variable	All Persons b	Non-South Asians b	Muslim b	Non-Muslim South Asians b	All South Asian b
Constant	100.33	99.13	97.32	95.63	99.06
OA Born in ineligible countries (% religion specific)	-0.95*	-0.30*	-1.10*	-2.48*	-1.70*
R-squared	.14	.01	.11	.05	.14

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. under-enumeration) 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 1 above), 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 3 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 under-estimated).

The resulting weighted figures for England and Wales are reported in table 3 above. As explained, the rates are all lower than the unweighted rates, especially for Muslims (91.5% after allowing for country of birth compared to 96.4 unweighted). The equivalent rate for other South Asians is just under 91%. The overall South Asian rate is slightly higher than either sub-group separately due to inclusion of a larger set of valid OAs (the Muslim and Non-Muslim rates for the 1150 OAs used in the calculation of All South Asian were 92.8% and 93.0% respectively). As explained above, the differences between these and the unweighted rates are due to the large stratification weight associated with the mainly the non-Asian areas which 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 following 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 not the result of lower registration rates of South Asians but an ecological fallacy. In other words it might have been due to lower registration rates

amongst the non South Asian population. The disaggregated analyses above dispelled this possibility. However, this does not mean that there were not geographical effects occurring whereby areas with larger South Asian populations experience lower registration amongst 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 disproportionate measurement error in stratum one.

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 & 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)

Valid OAs are shown in parantheses.

Table 5 breaks down the registration rate, comparing both with and without country of birth, of each group by the stratum in which they were sampled. As noted above, stratum one has the smallest proportion of South Asians (less than half a percent) and stratum five the largest (more than twenty percent). 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 five. 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 five, where South Asians make up more than 20% of the population is nearly 98%.

The pattern suggests two important conclusions. The first is substantive. The larger the South Asian community, the better mobilised and the more politically engaged they become. In other words, registration appears to be affected by belonging to a 'religious enclave' and therefore the potential mobilising affect of living in cohesive communities as. Conversely 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. The second implication is methodological. The South Asian rates for stratum one and two are sufficiently low to arouse suspicion that they may indeed be too low (as speculated above). In fact these estimates are likely to be unreliable as they are only based on a relatively small number of OAs. For example there are only two areas where the denominator is greater than 6.5 (not adjusted for country of birth) in stratum one for non-Muslim South Asians. However, given that the seemingly direct and consistent relationship between South Asian population concentration and registration, it seems entirely plausible that rates in these areas sparsely populated by South Asians do indeed have low registration rates for those communities.

This 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 (not shown here).

Figure 3. Muslim electorate and registration (by Output Area)
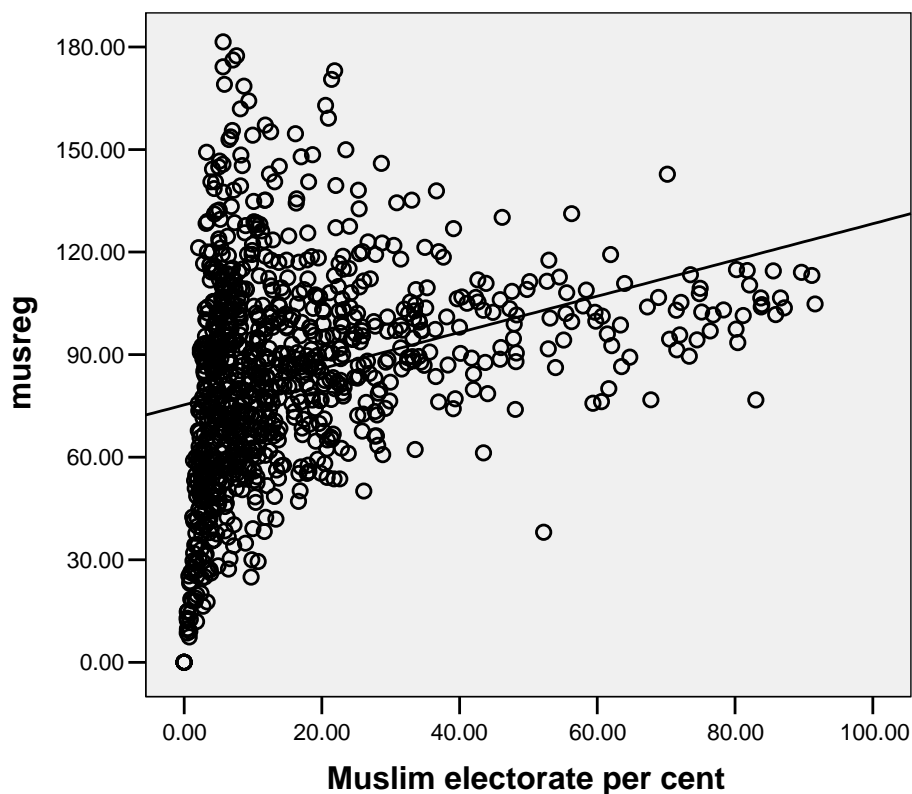
In order to confirm this we can confirm this relationship using a series of simple bi-variate OLS regression models (Table 6).[5] This confirms the positive relationship between Muslim registration and the number of Muslims living in an area. A similar statistically significant pattern exists for other 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 isn't affected by the proportion of non-South Asians in the area. Whether or not these relationships arise from mobilisation or social capital effects as suggested 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 & Sikh	-	-	-	+0.17*	-
OA % South Asian	-0.03*	+0.004	-	-	0.21*
R-squared	.01	.00	.09	.02	.05

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 one and two and hence trusting in the weighted national rates (rather than unweighted sample rates) reported in Table 3. In

---

[5]The apparent heteroscedasticity in Figure 3 is alleviated in the regressions as the regressions are weighted to reflect the number 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.

other words the national rate of registration for South Asians is approximately 86% but once country of birth is taken into account this rises to 93%.

So far we have demonstrated not only that registration rates are generally lower for South Asian communities, but 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 or by the socio-economic composition of the areas, rather than the religious composition. Again we use simple linear regression models of registration whilst controlling for ineligibility due to birthplace and a number of socio-economic indicators (see table 7 below).

Table 7. OLS Regression Models of OA registration (unadjusted for COB); Model 1 – Controlling for Country of Birth; Model 2 – 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
MODEL 1					
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 & Sikh	-	-	-	+0.21*	-
OA % South Asian	0.01	+0.04*	-	-	+0.06*
R-squared	.14	.02	.13	.08	.15
MODEL 2					
Constant	77.02	63.42	91.74	89.11	111.50
OA Born in ineligible countries (% religion specific)	-1.01*	-0.41*	-0.85*	-2.50*	-1.44*
OA % Muslim	-	-	+0.14*	-	-
OA % Hindu & Sikh	-	-	-	+0.03	-
OA % South Asian	+0.01	+0.04*	-	-	-0.03
Socio-economic variables					
OA Unemployment (% religion specific)	-0.17*	-	-	-0.67*	-
OA Owner occupation (% religion specific)	+0.05*	+0.03*	+0.16*	+0.14*	+0.10*
OA Manufacturing (% all persons)	-0.10*	-0.06*	-	+0.38*	-
OA Long Term Ill (% all persons)	-	-	-0.63*	-	-0.34*
OA Unemployment (% all persons)	-	-0.31*	-	-	-0.52*
OA Degree (% all persons)	-0.05*	-	-	-	
OA Pensioners (% all persons)	-	-	+0.19*	-	
OA High social class (% all persons)	-	-	-0.25*	-0.31*	-0.48*
Ward F-T students (% all persons)	+0.24*	+0.25*	-	-	-
Ward Non-migrants (% all persons)	+0.25*	+0.39*	-	-	-
R-squared	.20	.08	.18	.15	.19
Number of valid OAs	**1795**	**1795**	**936**	**753**	**1169**

* = Significant at the 95% 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.

The variables used in the model (table 7) include social and demographic variables

which measure characteristics for that religious group (Muslim or non-South Asian

Muslim, where available) and also characteristics of the area as a whole. Table 7 shows the results of the analyses.[6] 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^2). However, a number of interesting findings do emerge.

Model 1 (table 7) demonstrates that the positive association between Muslim registration and the geographical concentration of that 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 Muslims. For both Muslims and non-Muslim South Asians, the percent 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 sub-groups there is a significant negative effect related to the percent born outside eligible countries. There are smaller effects for non South Asians and the overall population. These results substantiate our earlier findings.

Model 2 in Table 7 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

---

[6] 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.

the unregistered. In fact the model shows that for every one percent increase in those born in ineligible countries there is just over a one percent decrease in registration. The model also confirms that overall registration isn't 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 and the number of students[7], and negatively correlated with unemployment and manufacturing.

The pattern for non-South Asians is fairly similar to the overall model, with a negative effect for the percent born in ineligible countries and a small positive effect for the percent South Asian. Significant control variables include the percentage non-migrants, students and owner occupier (all positively signed) and manufacturing and unemployment (negatively signed).

Table 7 also shows separate models for different religious groups. As noted above, we include variables which measure characteristics only for that group where available, and also characteristics of the area as a whole. For South Asians, some interesting patterns emerge. After controlling for socio-economic composition, whilst the proportion of the religion group born outside of eligible countries considerably

---

[7]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 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 F-T student variable is small and negative, suggesting the positive model coefficient is affected by correlation with other explanatory variables.

dampens all South Asian and both religion sub-group registration rates, there is no relationship between densely populated South Asian areas 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 sub-group registration rates significantly positively affected by home ownership for all groups, whilst long-term illness, unemployment and high social class status are generally negatively associated with South Asian registration.

Discussion

There are numerous ways of estimating registration rates, the implementation of which depends largely on the availability and quality of data. Whilst the simplest measures are based on a comparison of the voting age population and the number of registered electors, there are a number of refinements that can be made. In our analyses the number of electors (the numerator) excluded electors on the register who were known to have died or were yet to reach the age of 18. Our denominator or population count also takes into account the coming of age of attainers and estimated deaths. We also allow for the prospect of ineligibility due to nationality which is found to have a significant impact on registration rates, especially for Muslim communities. Using these data we find that, at the constituency level there is a negative association between the size of their South Asian population and the level of registration.

This is supported by OA level data disaggregated by religion derived from the electoral registers. These data show that South Asian adults are less likely to be registered than their non-Asian counterparts. However, this can be partly accounted for by differences in country of birth. In particular, Muslim communities have lower rates of registration than other South Asian communities before adjusting for ineligibility due to country of birth. After allowing for this, the national (weighted) registration rate for both Muslim and non-Muslim South Asians is approximately 93%. Indeed 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.

Perhaps equally important are our findings concerning the geography of registration for South Asian communities. In brief, in areas where South Asian populations are more concentrated, rates of registration for South Asian electors are much higher. However a negative correlation at the area level between registration for the population as a whole and the percentage of the population from South Asian communities means that this relationship is easily missed in aggregate analyses. Furthermore, studies based on particular types of area (e.g. exclusively areas of large Asian populations) may turn up misleading results. Indeed as we have shown, the opposing correlations for overall and South Asian registration made our estimation more difficult leading to differences between weighted and unweighted results, though this was resolved in favour of the weighted results for estimating national registration.

In particular 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 also observed for non-Muslim South Asians and all South Asians, though this did appear to reflect the socio-economic composition of areas. These results strongly suggest that South Asians living in South Asian communities are much more likely to be registered to vote than where they are geographically dispersed, a finding which is repeated for analysis of turnout (see Cutts and Fieldhouse, forthcoming, which also shows that on average turnout for South Asians is as high as, if not higher than, that for the rest of the population). This gives prima facie support to the argument that religious enclaves provides basis for political engagement through community networks and mobilisation (Le Lohe, 1998). To substantiate that requires further research and is beyond the scope of this paper. What we have focused on is providing a unique and nationally representative analysis of voter registration in 2001 disaggregated by religion, and with it evidence of lower rates of registration for South Asian electors particularly for those living outside the main areas of South Asian settlement.

References

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

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

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

Anwar, M. (1990) Ethnic Minorities and the Electoral Process: Some Recent Developments, in H. Goulbourne (ed.), *Black Politics in Britain*, UK: Gower Publishing Company

Cummins, C., Winter, H., Cheng, K. K., Maric, R., Silcocks, P. & Varghese (1999): An Assessment of the Nam Pehchan Computer Program for the identification of names of South Asian ethnic origin, *Journal of Public Health Medicine* 21(4), pp401-406

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 (2005) *Understanding electoral registration: The extent and nature of non-registration in Britain*: London, Electoral Commission

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

Fieldhouse, E and Cutts, D (2006) 'Voter Turnout in British South Asian Communities at the 2001 General Election' CCSR Working papers 2006 No 2 http://www.ccsr.ac.uk/publications/working/2006-02.pdf

Harding, S; Dews, H and Simpson, S (1999). 'The potential to identify South Asians using a computerised algorithm to classify names'. Population Trends 97 (Autumn): 46-49)

IDEA (2002) Voter Turnout Since 1945: A Global Report http//www.idea.int/vt/survey/voter_turnout8.cfm   accessed 5th June 2005

Johnston, R J and Pattie C J (1997) Towards an understanding of turnout at general elections: voluntary and involuntary abstentions in 1992. Parliamentary Affairs, 50, 1997, 280-291.

LeLohe, M.J. (1998) 'Ethnic minority participation and representation in the British electoral system' in S. Saggar (ed) *Race and British electoral politics*, London, UCL Press

LeLohe, M.J. (1990) The Asian Vote in a Northern City, in H. Goldbourne (ed.), *Black Politics in Britain*. U.K: Gower Publishing Company.

Lijphart, A. (1997) 'Unequal Participation; Democracy's Unresolved Dilemma' *American Political Science Review* 91 (March) 1-14

Martin D (2002) 'Output areas for 2001'. Chapter 3 in Rees P, Martin D and Williamson P (ed.) (2002). *The Census Data System*, John Wiley & Sons, Chichester, 37-46.

McDonald, M. & Popkin, S. (2001) The Myth of the Vanishing Voter, American Political Science Review, 24 (4): 963-74

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*, 23 (4): 278-85

Purdam, K; Fieldhouse E; Russell, A and Kalra V (2002) *Voter engagement among black and minority ethnic communities,* Electoral Commission. London

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* Winter 2005, 122: 27-34

Saggar, S (1998a) The General Election 1997: *Ethnic minorities and electoral politics*. London, Commission for Racial Equality

Saggar, S (1998b) (ed) *Race and British electoral politics* London, UCL Press

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

Todd JE and Butcher B (1981) Electoral Registration in 1981. London HMSO for OPCS.