



2009-10 Citizenship Survey

Technical Report



2009-10 Citizenship Survey
Technical Report

Ipsos MORI and TNS-BMRB

April 2011
Department for Communities and Local Government

© Queen's Printer and Controller of Her Majesty's Stationery Office, 2011

Copyright in the typographical arrangement rests with the Crown.

You may re-use this information (not including logos) free of charge in any format or medium, under the terms of the Open Government Licence. To view this licence, visit <http://www.nationalarchives.gov.uk/doc/open-government-licence/> or write to the Information Policy Team, The National Archives, Kew, London TW9 4DU, or e-mail: psi@nationalarchives.gsi.gov.uk.

This document/publication is also available on our website at www.communities.gov.uk

If you require this publication in an alternative format please email:
alternativeformats@communities.gsi.gov.uk

Any enquiries regarding this document/publication should be sent to us at:

Department for Communities and Local Government
Eland House
Bressenden Place
London
SW1E 5DU
Telephone: 030 3444 0000

April 2011

ISBN: 978-1-4098-2940-9

Acknowledgements

First and foremost our thanks go to all the respondents who gave up their time to take part in this survey.

Colleagues at Ipsos MORI and TNS-BMRB who have made a significant contribution to this project include Kathryn Gallop (Ipsos MORI), Kully Kaur-Ballagan (Ipsos MORI), Pamela Bremner (Ipsos MORI), Patten Smith (Ipsos MORI), Keith Bolling (TNS-BMRB), Tim Hanson (TNS-BMRB), Sam Sullivan (TNS-BMRB) and Rebecca Hamlyn (TNS-BMRB).

We would also like to thank those in the Operations Department and Data Delivery teams who have contributed to the project, in particular the many interviewers and wider Field Team who have worked on this survey.

Thanks also go to members of the Project Board and the Technical Advisory Group at the Department for Communities and Local Government (DCLG) who have contributed to the development of this project.

Finally, we would like to pay special thanks to Suzanne Cooper, Kate Millward, Reannan Rottier, Philippa Robinson, Elaine Wedlock, Farzana Bari, Helen Smith and their colleagues at DCLG.

Information on the Citizenship Survey and associated publications are available from:

<http://www.communities.gov.uk/communities/research/citizenshipsurvey/>

Suzanne Cooper
Department for Communities and Local Government
7/E8 Eland House
Bressenden Place
London
SW1E 5DU

Citizenship.survey@communities.gsi.gov.uk

Contents

Section 1	Introduction	6
Section 2	Sampling	8
	2.1 Overview	8
	2.2 Core sample	10
	2.3 Ethnic minority boost sample	14
	2.4 Muslim boost sample	17
	2.5 Sampling at addresses	20
Section 3	Questionnaire development and piloting	21
	3.1 Introduction	21
	3.2 Cognitive testing	21
	3.3 Dress rehearsal	25
	3.4 Questionnaire content	26
Section 4	Fieldwork	28
	4.1 Introduction	28
	4.2 Briefings	28
	4.3 Interviewer materials	29
	4.4 Contact procedures	30
	4.5 Screening procedures	32
	4.6 Screening and interviews with non-English speakers	36
	4.7 Maximising response	38
	4.8 Interview length	39
Section 5	Response rates	40
	5.1 Introduction	40
	5.2 Core sample	40
	5.3 Boost sample – focused enumeration screening	40
	5.4 Boost sample – ethnic minority direct screening	41
	5.5 Boost sample – Muslim direct screening	41

Section 6	Data processing	49
	6.1 Editing	49
	6.2 Coding	49
	6.3 Derived variables	50
	6.4 Data outputs	50
Section 7	Weighting	52
	7.1 Introduction	52
	7.2 Calculation of core sample weights	54
	7.3 Calculation of combined sample weights	63
Section 8	Standard errors	73
	8.1 Introduction	73
	8.2 Sources of error in surveys	73
	8.3 Standard errors for complex sample designs	74
	8.4 Design factor (deft)	74
	8.5 Confidence Intervals	75
	8.6 Standard errors for the 2009–10 Citizenship Survey	76
Section 9	Data User Guide	97
	9.1 Introduction	97
	9.2 Selecting cases for analysis	97
	9.3 Variables	98
	9.4 Multiple response questions	98
	9.5 Missing values	98
	9.6 Weighting	98
	9.7 Significance testing and standard errors	99
Section 10	References	100
Section 11	Annexes (produced under separate cover to this report)	101
	Annex A: Interviewer instructions	
	Annex B: Advance letters	
	Annex C: Leaflet	
	Annex D: Postcard	
	Annex E: Questionnaire	
	Annex F: Derived variables	
	Annex G: Quality Indicators for 2009–10 Citizenship Survey	

Section 1

Introduction

1. This report describes the methodology of the 2009–10 Citizenship Survey, also known as the Communities Study in the field. The study was carried out by Ipsos MORI in partnership with TNS-BMRB, on behalf of the Department for Communities and Local Government (DCLG). This report marks the sixth year of the study, and the third since a continuous design was introduced to allow for quarterly monitoring of key indicators. The data are published on a quarterly basis by the Department for Communities and Local Government.
2. The 2009–10 Citizenship Survey is a flagship survey for DCLG; used to measure performance as well as to inform and develop complex policy areas. The survey provides a wealth of information and data for a range of stakeholders across government and the wider research community. The objectives of the 2009–10 survey are as follows:
 - To continue to develop a robust, up-to-date and nationally representative evidence base on the key areas for which DCLG is responsible – including cohesion, racial and religious prejudice and discrimination, community empowerment, volunteering and civic participation – and to better understand and target policy in these areas;
 - To generate a thorough understanding of how attitudes and behaviours vary across sub-groups within the general population (including religion, age, sex, ethnicity, country of origin and sexual identity); and
 - To measure performance against a number of target indicators.
3. The importance of the survey data is highlighted by the fact that in March 2008 it was awarded National Statistics status. This means that the outputs have been certified by the UK Statistics Authority as compliant with the Code of Practice for Official Statistics.
4. The anonymised data sets for the Citizenship Survey are publicly available from the UK Data Archive (<http://www.data-archive.ac.uk/>), and quarterly statistical releases of the Citizenship Survey are available on the DCLG website:

<http://www.communities.gov.uk/communities/research/citizenshipsurvey/quarterlystatisticalreleases/>

5. The Survey includes questions related to a number of topics, including: identity and social networks, local community, influencing political decisions and local affairs, volunteering, objective empowerment, race, religion, and rights and responsibilities. It also collects socio-demographic data about respondents. The findings from the 2009–10 study are presented in topic reports and published on the DCLG website.
6. In line with the 2008-09 survey, the 2009–10 Citizenship Survey delivers a nationally representative sample of the adult population (aged 16 years and over) of England and Wales. The sample contains approximately 10,000 interviews achieved over the course of the year (c.2,500 each quarter). In addition, there are 5,000 boost interviews with ethnic minorities. For the first time, the study also contains a Muslim boost, which, when combined with Muslim respondents from the main sample and ethnic minority boost, generates over 3,000 interviews with Muslim respondents.

Section 2

Sampling

2.1 Overview

2.1.1 The 2009-10 Citizenship Survey sample comprised three main elements:

- a core sample designed to deliver 10,000 interviews per year with a random sample of the adult (aged 16 years and over) population living in private households in England and Wales;
- an ethnic minority boost sample designed to deliver approximately 5,000 interviews per year with adults living in private households and belonging to eligible ethnic minority groups¹ in England and Wales; and
- a Muslim boost sample designed to ensure that in total (i.e. from all sample components) 3,000 interviews were conducted per year with Muslim adults living in private households in England and Wales.

2.1.2 The ethnic minority boost sample comprised two distinct sub-samples:

- a *focused enumeration boost sample* in areas where the concentration of the (all-age) ethnic minority population was greater than 1 per cent and under 18 per cent; and
- a *high concentration direct boost sample* in areas where the ethnic minority population concentration was 18 per cent or more.

2.1.3 Ethnic minority individuals were identified by means of focused enumeration, i.e. proxy screening, and direct screening respectively in these two sub-samples.

¹ eligible ethnic minority groups were:

- Mixed: White and Black Caribbean
- Mixed: White and Black African
- Mixed: White and Asian
- Mixed: Other Mixed
- Asian or Asian British: Indian
- Asian or Asian British: Pakistani
- Asian or Asian British: Bangladeshi
- Asian or Asian British: Other Asian
- Black or Black British: Black Caribbean
- Black or Black British: Black African
- Black or Black British: Other Black
- Chinese or Other Ethnic Group: Chinese
- Chinese or Other Ethnic Group: Other Ethnic Group

Focused enumeration was undertaken in core sample primary sample units around issued core sample addresses. In the high concentration boost (18%+) areas, fieldwork was conducted in primary sample units which were selected independently of core sample primary sample units. Eligible sample members were identified through direct screening.

2.1.4 The Muslim boost sample comprised two elements:

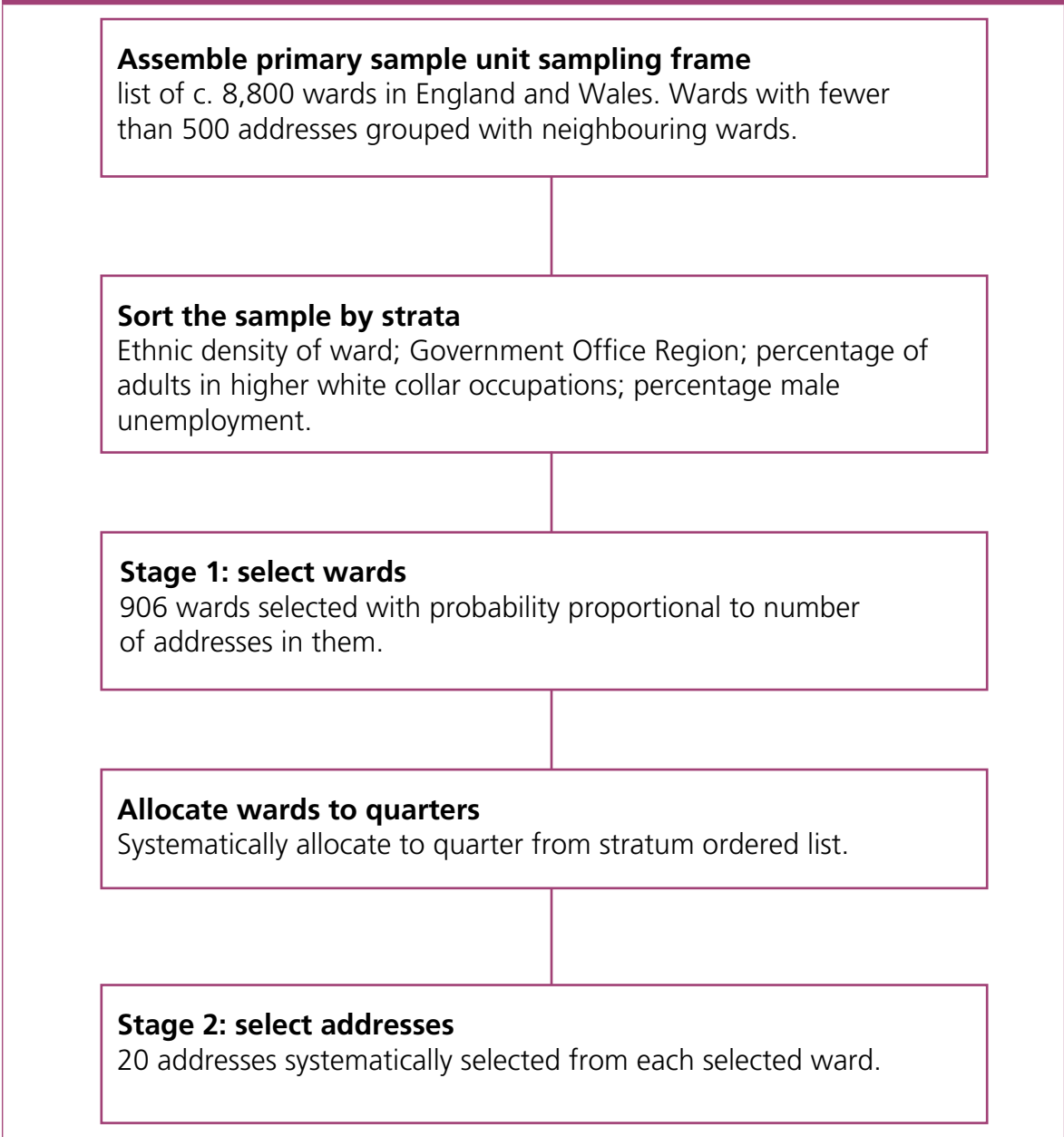
- addresses issued as part of the high concentration ethnic minority boost sample where no eligible individuals were identified were then screened for the presence of Muslims; and
- a separate sample of addresses was drawn for screening in areas in which at least 2.5 per cent of the population was Muslim. These addresses were screened for the presence of Muslims by means of a face-to-face screening interview.

2.1.5 The Citizenship Survey used a two-stage stratified random sample design to obtain addresses. In the first stage, for both core and ethnic minority boost samples, a systematic sample of (grouped²) wards was selected. In the second stage, addresses were systematically sampled within the selected wards from the small-user postcode address file. This approach was also used in previous sweeps of the survey. A similar approach was taken in selecting addresses for the Muslim boost sample; however, (grouped) output areas were used as primary sample units instead of wards. This process is shown (for the core sample) in Figure 2.1.

2.1.6 There are about 8,800 wards in England and Wales, each containing on average around 2,500 addresses. Using wards to cluster the sample gives a balance between having manageable interviewer workloads within a controlled geographical area and the loss of statistical efficiency resulting from a clustered sample.

² Wards containing fewer than 500 addresses were grouped with neighbouring wards to ensure selected addresses were not too close together.

Figure 2.1: Two-stage stratified sampling of core sample addresses



2.2 Core sample

2.2.1 Table 2.1 summarises the assumptions underlying the core sample design. It was expected that an issued sample of 18,120 addresses would deliver the desired achieved sample of 10,000 interviews.

Table 2.1: Assumptions underlying core sample design

Number of primary sample units	906
Addresses issued in each primary sample unit	20
Total addresses issued	18,120
Deadwood rate	8%
Response rate	60%
Achieved sample	10,002

- 2.2.2 A list of all wards in England and Wales, including a count of the number of postcode address file addresses³ in each, was generated. Wards containing fewer than 500 addresses were combined with neighbouring wards to form primary sample units in order to ensure that all interview addresses were reasonably well separated.
- 2.2.3 Before selection, primary sample units were sorted into three strata based upon the proportion of the total population in each ward in an eligible ethnic minority group:
- high concentration stratum in which 18 per cent+ of the population (according to the 2001 Census) were in an eligible ethnic minority group;⁴
 - medium concentration stratum in which at least one per cent, but less than 18 per cent of the population were in an eligible ethnic minority group; and
 - low concentration stratum in which less than one per cent of the population were in an eligible ethnic minority group.
- 2.2.4 In the 2008–09 survey the high concentration stratum was divided into two: a high density stratum in which 18 per cent+, but less than 60 per cent, of the population were in an eligible group and a super-high density stratum in which 60 per cent+ of the population were eligible. This was done because in 2008–09 different primary sample units sampling rates were applied to high density and super-high density primary sampling units when selecting the ethnic minority boost sample. Changes to the ethnic minority boost sample design in 2009–10 obviated the need for subdividing the high concentration stratum (see paragraph 2.3.4 to 2.3.6).

³ More exactly, the number of delivery points in the small user version of the postcode address file.

⁴ The 18 per cent+ threshold was used in previous rounds of the survey. It was originally deemed an appropriate level at which to separate the strata in which face-to-face screening was carried out from the strata in which screening was by means of focused enumeration, bearing in mind the need to balance survey costs against sample quality in the interests of maximising value for money.

- 2.2.5 Within each of these strata, primary sample units were then sorted by Government Office Region.⁵ Within each resulting Government Office Region-within-ethnic minority concentration stratum, the primary sample units were listed by proportion of adult population in higher white collar occupations and then divided into three approximately equal groups (equal in terms of address count rather than number of primary sample units). Within each of these strata, the primary sample units were then sorted by proportion of males aged 16–74 who were unemployed.
- 2.2.6 Proportionate stratification generally improves the precision of survey estimates for variables which correlate with the variables used for stratification, the stratification variables were chosen on the basis of their likely association with a range of survey variables. The stratifiers used were closely based on those used in the 2008–09 survey.⁶ The 2008–09 sample was stratified by proportion of the population in non-manual occupations within Government Office Region whereas the 2009–10 sample was stratified by proportion of population in higher non-manual occupations. It is highly likely that they are very similar and will undoubtedly have very similar effects on the precision of survey estimates. The way in which the list of primary sample units was ordered before selection is illustrated in Figure 2.2.

⁵ See Office for National Statistics, *Geography – Government Office Regions*. Accessible at: <http://www.statistics.gov.uk/geography/gor.asp> for more details

⁶ See NatCen, (2008), *2008-9 Citizenship Survey: Technical Report*, (London: Communities and Local Government).

- 2.2.7 A total of 906 primary sample units were selected with probability proportional to postcode address file address count from this list, using the method of random start and fixed interval. Twenty addresses were then selected systematically from each sampled primary sample unit.
- 2.2.8 This delivered an equal probability sample of addresses that was proportionately stratified by:
- ethnic minority concentration of address’s primary sample unit
 - Government Office Region;
 - concentration of workers in higher white collar occupations in address’s primary sample unit, and
 - percentage of working age males that are unemployed in each address’s primary sample unit.
- 2.2.9 Core sample primary sample units were allocated systematically to quarters within the survey year in order to preserve the sample stratification. This was done by determining a random sequence of quarters and then allocating sampled primary sample units sequentially to these quarters. For example if the random sequence of quarters had been: Q3, Q2, Q4, Q1, then the sampled primary sample units would be allocated to quarters as shown in Table 2.2. This approach ensured that each quarterly sample of primary sample units was, as far as possible, stratified by the same variables as the annual sample.

Table 2.2: Example of random allocation of PSUs for quarters

Order of initial selection	Allocated quarter
1	3
2	2
3	4
4	1
5	3
6	2
(sequence continued).....	(sequence continued).....

2.3 Ethnic minority boost sample

- 2.3.1 Two methods were used to draw the ethnic minority boost sample: focused enumeration and direct screening. These methods are described below and summarised in Table 2.3.

Focused enumeration

2.3.2 For each core sample address issued in the 622 primary sample units, where more than 1 per cent, but less than 18 per cent, of the population belonged to an ethnic minority group (i.e. primary sample units in the medium concentration stratum defined above), the two preceding and two following addresses on the postcode address file were drawn for focused enumeration screening. This meant that 80 addresses were issued for focused enumeration in each primary screening unit. In all, 12,440 core sample and 49,760 focused enumeration addresses were issued in the medium concentration stratum. It was estimated that focused enumeration addresses would deliver around 550 interviews with members of eligible ethnic minority groups.

2.3.3 In summary, focused enumeration involved the following steps:

1. Core sample respondents were asked during screening if anybody of an eligible ethnic minority group lived at each of the four pre-selected adjacent focused enumeration sample addresses.
2. If it was reported that *no eligible person lived at a neighbouring address*, no attempt was made to contact it.
3. In any other case (if eligible respondents were identified or respondents did not know the eligibility of neighbouring addresses), the interviewer visited the address in order to conduct a face-to-face screening interview designed to identify residents belonging to eligible ethnic minority groups.

Direct screening

2.3.4 The high concentration ethnic minority boost sample was drawn from primary sample units in which at least 18 per cent of the population belonged to an ethnic minority group. The sample was designed to minimise cluster size variation and involved sampling primary sample units with probability proportional to a synthetic size measure. The procedures are summarised below:

1. A size measure, s (essentially the estimated number of households (HH) containing at least one adult belonging to an eligible ethnic minority group) was calculated for each primary sample unit:

$s = plc$, where:

p = number of postcode address file addresses in the primary sampling unit

c = estimated proportion of households containing ethnic minority adults in the PSU (estimation of c required application of Labour Force Survey-estimated multipliers to primary sampling until level 2001 Census ethnic group population figures)

l = estimated address occupation rate (fixed at 92%).

2. Primary sample units were ordered by proportion of the population in eligible ethnic minority groups.
3. 292 primary sample units were systematically selected with probability proportional to s .
4. Within each sampled primary sample unit the number of addresses to select was determined according to the formula:

$$a = 15 / (l c r_s r_m), \text{ where:}$$

$$r_s = \text{estimated screening response rate (fixed at 85\%)}$$

$$r_m = \text{estimated main interview response rate (fixed at 52.5\%).}$$

5. Addresses were selected systematically from the postcode ordered postcode address file.

2.3.5 A total of 30,613 addresses were selected in this way. These were predicted to deliver 4,500 interviews with members of eligible ethnic minority groups and 190 additional interviews with Muslims belonging to ethnic minority groups that are not eligible for inclusion in the ethnic minority sample. The above design ensured that all screened addresses were sampled with equal probability. If response rate and eligibility assumptions proved to be correct, the design would also have delivered exactly 15 interviews with members of eligible ethnic minority groups in each primary sample unit.

2.3.6 The selection procedures described above differed from those used in 2008-09 (although they are very close to those used in the original Citizenship Survey in 2001). The 2008-09 survey managed cluster size by issuing smaller numbers of addresses for screening in the super-high density stratum primary sample units (primary sample units in which more than 60 per cent of the population was from an ethnic minority group), and adjusting the relevant primary sample unit selection probabilities to ensure that addresses were sampled with equal probability. Theoretically at least, the approach used in 2009-10 should lead to smaller cluster size variability than the approach used in 2008-09.

2.3.7 All addresses in the issued sample were visited by interviewers and screening interviews were attempted. The screening interview set out to identify all household members who belonged to an eligible ethnic minority group or who was a Muslim belonging to an ineligible ethnic group. Table 2.3 summarises the design for the ethnic minority boost sample, including the number of primary sample units, addresses issued and the estimated number of interviews.

Table 2.3: Summary of ethnic minority boost sample design

	Focused enumeration	Face-to-face screening	Total
Number of primary sample units	622	292	914
Total addresses issued	49,760	30,613	80,373
Expected yield rate	1.1%	14.7%	
Estimated achieved	550	4,500	5,050
Achieved interviews	740	4,540	5,280

2.3.8 Ethnic minority boost sample primary sample units were allocated systematically to quarters within the survey year in order to preserve the sample stratification, as described for the core sample in section 2.2.9.

2.4 Muslim boost sample

- 2.4.1 We estimated that the above core and ethnic minority boost samples would deliver 1,800 interviews with Muslims. The separate Muslim boost sample was therefore designed to deliver 1,200 additional interviews.
- 2.4.2 Preliminary examination of 2001 Census data demonstrated that for a fixed fieldwork budget and achieved sample size a higher coverage Muslim sample would be achieved if output areas were used as primary sample units instead of wards. The Muslim boost sample was therefore selected by means of a two-stage design in which output areas rather than wards were used as primary sample units. The sample was drawn from output areas in which at least 2.5 per cent of the population was identified as being Muslim in the 2001 Census. Output areas containing fewer than 125 postcode address file addresses satisfying this criterion were attached to geographically close eligible output areas to construct a population file of Muslim boost sample primary sample units.
- 2.4.3 A disproportionately stratified sample design was used in order to limit the number of addresses issued for screening in primary sample units where Muslims were less highly concentrated according to the 2001 Census. To this end primary sample units in the population output area file were divided into two strata to identify eligible respondents within different types of areas:
- those in which we estimated that 10 per cent or more households contained 1+ Muslim adults; and
 - those in which we estimated that fewer than 10 per cent of households contained 1+ Muslim adults.

- 2.4.4 In the 10 per cent+ stratum, a synthetic size measure, similar to that used to select the high concentration ethnic minority sample, was constructed and used to select primary sample units with probability proportional to size. Addresses were then selected with probability inversely proportional to this size measure, thereby ensuring that addresses were sampled with equal probability within the stratum.
- 2.4.5 In the under 10 per cent stratum, a simple equal probability sample of primary sample units was selected using the standard method of random start and fixed interval, and then all addresses were sampled within sampled primary sample units, again ensuring that addresses were sampled with equal probability within the stratum. The overall address selection probability for addresses in the under 10 per cent stratum was one quarter of that used in the 10 per cent+ stratum in order to control the total number of addresses requiring screening.

The 10 per cent+ stratum

- 2.4.6 The procedures used to select primary sample units and addresses in the 10 per cent+ stratum are described below.

1. A size measure, s (essentially the estimated number of HHs containing at least one Muslim adult) was calculated for each primary sample unit:

$s = plc$, where:

p = number of postcode address file addresses in the primary sampling unit;

c = estimated proportion of households containing Muslim adults in the primary sampling unit (estimation of c required application of LFS-estimated multipliers to primary sampling unit level 2001 Census ethnic group population figures);

l = estimated address occupation rate (fixed at 92%).

2. Primary sample units were ordered by proportion of the population that was Muslim.
3. 222 primary sample units were selected systematically with probability proportional to s .
4. Within each sampled primary sample unit the number of addresses to select was determined according to the formula:

$a = 5/(lcr_s r_m)$, where:

r_s = estimated screening response rate (fixed at 85%)

r_m = estimated main interview response rate (fixed at 50%).

5. Addresses were selected systematically from the postcode ordered postcode address file.

2.4.7 The above design ensured that all screened addresses were sampled with equal probability. If response rate and eligibility assumptions proved to be correct the design would also have delivered exactly five interviews with Muslim individuals in each primary sample unit.

Under 10 per cent stratum

2.4.8 Forty-four primary sample units were selected systematically with equal probability, and all addresses in each primary sample unit were selected for screening.

2.4.9 The numbers of issued addresses and our estimates of the number of achieved interviews in each of the two Muslim boost strata are shown in Table 2.4 below.

Table 2.4: Summary of Muslim boost sample design

	Number of primary sampling units	Issued addresses	Achieved interviews with Muslims (estimated)	Achieved interviews with Muslims (actual)
10 per cent+ stratum	222	11,696	1,110	1,402
Under 10 per cent stratum	44	8,073	147	153
Total	266	19,769	1,257	1,555

2.4.10 Muslim boost sample primary sample units were allocated systematically to quarters within the survey year in order to preserve the sample stratification, in the manner described for the core sample at section 2.2.9.

2.5 Sampling at addresses

- 2.5.1 A small proportion (under 2%) of addresses sampled from the postcode address file contained more than one dwelling unit.⁷ At these addresses interviewers selected one dwelling unit using the Kish grid procedure.⁸
- 2.5.2 At selected dwelling units, one adult aged 16 or over was selected from all eligible⁹ adults in the dwelling unit using the Kish grid procedure.

⁷ A dwelling unit is a unit within a postcode address file address which has a separate front door and is accessible only to the subset of postcode address file address residents living behind that front door.

⁸ A Kish grid is a technique used where interviewers who have been issued with a sample of household addresses can sample individuals on the doorstep. It involves using a random selection technique, and is devised to ensure all individuals eligible in the household have an equal chance of selection. See L. Kish (1949), 'A Procedure for Objective Respondent Selection Within the Household', *Journal of the American Statistical Association*.

⁹ Eligibility was defined as follows:

- core sample: aged 16 or over
- ethnic minority boost sample: aged 16 or over and belonging to an eligible ethnic group
- Muslim boost sample: aged 16 or over and Muslim.

Section 3

Questionnaire development and piloting

3.1 Introduction

- 3.1.1 The 2009–10 questionnaire was largely based on the 2008–09 Citizenship Survey. However, development work did generate a small number of changes. Questions were added or removed, where relevant, to reflect changing policy priorities and developments in the objectives of the survey since the 2008-09 study.
- 3.1.2 Development of the questionnaire had two main objectives:
- to test proposed new questions; and
 - to test the questionnaire as part of a pilot study (referred to here as the dress rehearsal) before the main fieldwork.

3.2 Cognitive testing

- 3.2.1 The first objective of testing proposed new questions was achieved through two phases of cognitive testing.
- 3.2.2 Cognitive interviews are used as a method of question testing to explore the cognitive processes involved when people interpret and respond to survey questions. Cognitive interviews are qualitative in nature, involving a small sample and in-depth probing techniques. They help to reduce measurement error by ensuring questions are designed so that respondents understand and are willing and able to answer them. The techniques used help us to establish how questions are understood by respondents, how they arrive at their responses, how confident they are in their answers, and to identify any problems that questions may pose. In particular, cognitive interviews can explore reasons for any problems and respondent reactions to questions that may be of a sensitive nature.

- 3.2.3 For the 2009–10 survey, the main aims of the cognitive pilot were to explore respondents’:
- understanding of terms used in new questions, e.g. ‘place of worship’, ‘immigrant’, and ‘violent extremism’;
 - ability to understand and interpret new showcards, identifying, for example, any problems with long showcards;
 - ability to answer in relation to set timeframes and areas, e.g. ‘a year ago’, ‘within half a mile of your home’;
 - judgements and interpretations, e.g. what constitutes ‘a significant drop in income’; and
 - reactions to new potentially sensitive questions, e.g. when asked about their responses to materials promoting violent extremism.
- 3.2.4 Interviews lasted on average between 30 and 45 minutes. Researchers made detailed notes during the interview, which were collated following completion of the cognitive testing. The interview notes were used to produce findings and recommendations on changes to questions.
- 3.2.5 All cognitive interviews were conducted by researchers from Ipsos MORI and TNS-BMRB.
- 3.2.6 In Round 1, interviews were carried out with members of the public in the following six locations: Birmingham, Manchester, Hull, Bromley, Oxford, and West London. These areas were chosen to ensure a broad geographic spread. Interviewers were assigned addresses within these areas, and recruited on a face-to-face basis to meet the quotas listed below. In total, 20 interviews were conducted, with each respondent receiving £20 for taking part. Interviewers were given quotas of respondent characteristics to ensure that the sample included respondents who:
- were from a range of age groups;
 - were of different sexes;
 - lived in a range of urban and rural areas;
 - were of different ethnic backgrounds; and
 - were of different faiths, including Muslims.

- 3.2.7 The interviews were carried out in December 2008 and tested the following question areas.
- question 'Gtm2': Whether the respondent would mind having certain facilities near their home (specifically a primary school, a nightclub, a Gypsy or Irish Traveller site);
 - the respondent's feelings (positive or negative) towards different groups (specifically women, people under 30, Muslims);
 - the extent to which the respondent trusts their local councillor;
 - (If religious) how often the respondent has visited a place of worship in the last 12 months;
 - questions 'Letin/Redgps': The respondent's views on immigration in the UK (i.e. whether it should be increased, reduced or remain the same);
 - questions 'Skiqua/Hapski': Whether the respondent is doing a job that fully uses their skills/qualifications, and if not, how they feel about this; and
 - questions 'Finhap/Finrec': How, if at all, the respondent has been affected by the economic downturn and how they think their financial circumstances might change in the future.
- 3.2.8 Feedback from these interviews was used to refine and revise the tested questions. Following these revisions, some of the questions were deemed ready and suitable to be included in the 2009–10 study. In some cases, however, it was felt appropriate to conduct further testing – these questions were carried forward to the second round of cognitive interviewing.
- 3.2.9 In Round 2, interviews were carried out with members of the public in the following six locations: Birmingham, London, Leeds, Leicester, Reading and Luton. These areas were chosen to ensure a broad geographic spread. The recruitment and interview process was the same as for Round 1, with 21 interviews being conducted in total.
- 3.2.10 The interviews were carried out in February 2009 and tested the following question areas:
- questions 'Letin/Redgps': The respondent's views on immigration into the UK (i.e. whether it should be increased, reduced or remain the same);
 - questions 'Skiqua/Hapski': Whether the respondent is doing a job that fully uses their skills/qualifications, and if not, how they feel about this;
 - questions 'Finhap/Finrec': How, if at all, the respondent has been affected by the economic downturn and how they think their financial circumstances might change in the future;

- whether the respondent has had any exposure to/contact with violent extremism and, if so, their opinion of what they saw or heard;
- whether the respondent is aware of any groups in their area that support violent extremism, and what action they would take if they become aware of any such group; and
- the respondent's level of agreement with statements relating to Islam.

3.2.11 Following the testing of questions in both Round 1 and Round 2, and based on analysis of the taped recordings of the interviews, and interviewer notes, a set of findings and recommendations was produced.

3.2.12 Of the questions that were tested during cognitive interviewing, the following were included in the final questionnaire:

- Gtm2 (Whether mind having certain facilities near to home);
- Letin/Redgps (Whether think immigration should be increased, reduced, or remain the same) – included as an unprompted question with a pre-coded response list;
- Skiqua/Hapski (Whether job uses skills/qualifications and, if not, how respondent feels about this) – asked only of respondents with qualifications;
- Finhap/Finrec (Affect of economic downturn on respondent) – included as a prompted question;
- exposure to and opinions of violent extremism;
- respondent's awareness of any violent extremist groups in their area, and action respondent might take in response to this; and
- level of agreement with statements relating to Islam.

3.2.13 In addition, several questions that appeared in the 2008-09 questionnaire were removed. These questions covered the following topics:

- respondent's satisfaction with their home;
- being kept informed about services and benefits provided by local councils;
- motivation for involvement in volunteering;
- motivation for involvement in community activities; and
- how often respondent visits place of worship.

3.3 Dress rehearsal

- 3.3.1 The purpose of the dress rehearsal was to test the main survey procedures that would be used during mainstage fieldwork, test the CAPI¹⁰ program and check the length of the interview. Fieldwork took place between 28 January and 19 February 2009.
- 3.3.2 Interviews were carried out in 20 wards, purposively sampled to differ according to the density of ethnic minority or Muslim residents, so that there would be appropriate high- and low-density areas to test the various screening procedures. The mix of wards selected, and types of samples included, were:
- four low- or medium-density ethnic minority (core sample interviews);
 - six medium-density ethnic minority (core sample interviews and focused enumeration screening);
 - six high-density ethnic minority (ethnic minority + Muslim direct screening); and
 - four high-density Muslim (Muslim direct screening).
- 3.3.3 The wards were geographically spread across England and Wales, with interviewing taking place in both rural and urban areas.
- 3.3.4 In each ward where core interviewing took place, 30 addresses were randomly selected. In the core and focused enumeration wards, the two addresses either side of the core address were selected for focused enumeration screening. The number of addresses selected in wards where boost screening would take place differed depending on the density of ethnic minority/Muslim residents. A sufficient number of addresses were selected in each ward to achieve the target number of interviews (eight for ethnic minority and Muslim screening assignments; five for Muslim screening assignments).
- 3.3.5 Twenty interviewers attended a full-day face-to-face briefing with researchers and members of DCLG. The design of the briefing and the survey instructions were intended to form the basis for these elements of the mainstage survey. Similarly, the survey materials, such as advance letters, leaflets and address contact sheets were those intended for use in main fieldwork. The procedures used to screen and select respondents at households were also those to be used at the mainstage. The dress rehearsal was a full and comprehensive test of all procedures and materials.

¹⁰ Computer Aided Personal Interviewing (CAPI) – interviews were conducted via a laptop using Quancept software. The program ensured that questions appeared on screen in a specific order, and that each respondent was asked the correct questions based on the answers they gave.

- 3.3.6 Initial contact with core respondents was made using an advance letter about the study. For boost screening and focused enumeration cases, interviewers were provided with a letter to give to potentially eligible respondents on the doorstep.
- 3.3.7 In total, 188 interviews were achieved. Interviewers were given a feedback form to complete and attended a full debrief with researchers and members of DCLG. As a result of feedback from interviewers, a number of amendments were made to various aspects of the survey:
- the contact sheets were shortened and simplified to make them easier to use
 - the content of the advance letter and leaflet were slightly revised;
 - a postcard was introduced. This gave a few basic details about the survey without giving too much background. It was intended that the postcards could be used for screening samples (where an advance letter was not delivered) to ensure interviewers were not always 'cold calling' at addresses;
 - a few changes were made to the briefing slides for main stage interviewers; and
 - the 'Top Tips' document was updated with feedback from interviewers (particularly on tips for screening at addresses and ways to introduce the survey).

3.4 Questionnaire content

- 3.4.1 The CAPI questionnaire used for the mainstage in 2009-10 consisted of the following modules:
- **Household composition** – details of people living in the selected households; identification of the household reference person; basic employment details of the respondent.
 - **Identity and social networks** – how the respondent perceives their national identity; basic demographic details about the respondent's friends.
 - **Your community** – the respondent's sense of belonging to, and views about, area of residence and other residents; how worried the respondent is about various types of crime; how the respondent does/would feel about living near various types of facility; respondent's satisfaction with local services.
 - **Influencing political decisions and local affairs** – involvement in local affairs; degree to which the respondent can affect political decisions at various levels; trust in institutions.

- **Volunteering and charitable giving** – the respondent’s involvement with organised groups; giving help through groups (formal volunteering); volunteering through an employer; opportunities for, and barriers to, formal volunteering; informal volunteering; charitable giving.
- **Objective empowerment** – the respondent’s involvement in community decision-making through formal roles and groups.
- **Race and immigration** – the respondent’s perceptions of racial prejudice, and of racial discrimination by public service organisations; experiences of discrimination in employment; their views on current levels of immigration.
- **Religion** – the respondent’s perceptions of religious prejudice; whether they personally have been subjected to religious discrimination; details of their religion and religious background.
- **Mixing** – whether the respondent mixes socially with people from different ethnic and religious groups to themselves; the respondent’s views on what would encourage people to mix more.
- **Respect** – whether the respondent feels they are treated with respect in various everyday situations.
- **Self identity** – factors that are important to the respondent’s sense of who they are.
- **Violent extremism** – whether the respondent feels it is right or wrong to use violent extremism as a form of protest; whether they personally have done anything to reduce support for violent extremism in the name of religion; whether they are aware of/been exposed to groups or materials promoting violent extremism in the name of religion.
- **Media usage** – information about media exposure.
- **Economic downturn and demographics** – whether the respondent has been affected in any way by the economic downturn; and whether the respondent has an illness or disability; tenure status; sexual identity; employment details of the household reference person (if they are not the respondent); employment details of the respondent; educational qualifications; income of respondent (and their partner).

3.4.2 A paper version of the questionnaire can be found at Annex E (under separate cover to this report).

Section 4

Fieldwork

4.1 Introduction

4.1.1 Fieldwork for the 2009–10 Citizenship Survey was conducted in four quarters between 1 April 2009 and 31 March 2010. All fieldwork was carried out by trained interviewers who are members either of Ipsos MORI’s field team or Kantar’s field team. Kantar operates on behalf of TNS-BMRB.

4.1.2 Fieldwork dates for 2009–10 were as follows:

- Quarter 1: 1 April 2009–30 June 2009. First reissue date – 8 May.
- Quarter 2: 1 July 2009–30 September 2009. First reissue date – 31 July.
- Quarter 3: 1 October 2009–31 December 2009. First reissue date – 3 November.
- Quarter 4: 1 January 2010–31 March 2010. First reissue date – 4 February.

4.2 Briefings

4.2.1 All interviewers working on the study received a one-day comprehensive briefing delivered by the research teams and field management staff of Ipsos MORI and TNS-BMRB. Whenever possible, a representative from the DCLG Citizenship Survey team also participated in the briefings.

4.2.2 In total, 46 briefings took place, attended by a total of 581 interviewers.

4.2.3 Each briefing included the following elements:

- An introduction to the survey: aims and objectives of the survey, information about DCLG and its ministers, an overview of the topics included in the questionnaire and examples of how study results are used;
- Fieldwork procedures: an explanation of the different sample types; procedures for the focused enumeration sample; non-contact and contact approaches; screening procedures; how to select dwelling units and

respondents; how to identify translation requirements and conduct survey language interviews;

- Survey materials: information and advice was given on using these before and during the interviews;
- Contact sheet exercises and discussions on maximising response rates;
- The interview: introducing the questionnaire and a practice session using the CAPI program to interview a dummy respondent, including use of showcards and shuffle cards; and
- Field administration: advice on carrying out fieldwork; key fieldwork dates; and field contact procedures.

4.3 Interviewer materials

4.3.1 In advance of each fieldwork quarter, interviewers received survey packs containing all the materials they needed for fieldwork in that quarter. The survey pack included:

- **Contact sheets** for each address, according to their sample type. Either core, core plus focused enumeration, focused enumeration, ethnic minority and/or Muslim;
- **Letters** according to their sample type. Either core letters for all core addresses or screening letters for all other sample types;
- **Leaflets** (glossy, full colour) providing an overview of the survey, with an example of survey results from a previous study year and contact details for either Ipsos MORI or TNS-BMRB;
- **Postcards** designed to be left at screening addresses, informing residents that a survey is being conducted in the area and providing contact details of either Ipsos MORI or TNS-BMRB. Interviewers had the opportunity to record their name on the reverse of the postcard;
- **Calling cards** for interviewers to leave their contact details;
- A **language showcard** (A4, laminated) with a short paragraph in each of the survey languages explaining why the interviewer was at their household and used to identify survey language needs when there was no common language between the interviewer and the household;
- **Faith and ethnicity screening showcards** (A4, laminated) used to identify eligible households and individuals at boost addresses. For interviewers with survey language skills, these showcards were also provided in relevant languages;

- A **violent extremism definition showcard** (A4, laminated) for use during the interview. For interviewers with survey language skills, this showcard was also provided in relevant languages;
- **Questionnaire showcards** (A4, bound with spiral plastic binding) for use during the interview. For interviewers with survey language skills, these showcards were also provided in relevant languages;
- A set of **shufflecards** for use during the volunteering section of the questionnaire. For interviewers with survey language skills, these shuffle cards were also provided in relevant languages;
- A paper version of **the questionnaire** as a preparation tool for the interviewer. For interviewers with survey language skills, this was also provided in relevant languages;
- **Interviewer instructions**, as a comprehensive guide to the study; and
- A **Top Tips** document, outlining advice and guidance for dealing with common respondent questions, and techniques for helping interviewers achieve interviews.

4.4 Contact procedures

Letters

- 4.4.1 All addresses in the core sample received a letter in advance of the interviewer's first visit explaining the purpose of the survey, how the address had been selected and stating that an interviewer would be calling at the address. The letter explained that all information would be confidential and stressed the importance of participation in the study. It also provided a named contact at either Ipsos MORI or TNS-BMRB, with an email address and telephone number, should selected households want further information about the study. The letter, and all communications with respondents, referred to the research as 'the Communities Study'.
- 4.4.2 Screening letters and postcards were designed for use with the boost sample addresses (focused enumeration, ethnic minority and Muslim boosts). These were left at addresses where no contact had been made at the first call, to make respondents aware of the study and that an interviewer had called at the address. Although the content was very similar to the core sample letter, it made clear the fact that not all addresses in this sample would contain eligible people for interview.
- 4.4.3 Copies of the core and screening letters are available in Annex B (under separate cover to this report).

Leaflets

- 4.4.4 All core addresses received a survey leaflet with their advance letter. This provided further information about the survey, including a chart showing results from previous surveys in the series. It also addressed potential concerns about data protection and provided contact details for Ipsos MORI or TNS-BMRB – a Helpline telephone number and email address – and website addresses for DCLG and the ‘Communities Study’.
- 4.4.5 Interviewers were given copies of the leaflets which could be issued to screening respondents.
- 4.4.6 A copy of the leaflet is available in Annex C (under separate cover to this report).

Translations

- 4.4.7 Core addresses in Wales received advance letters and leaflets in Welsh and English.
- 4.4.8 In addition to English, the survey could also be conducted using bilingual interviewers or interpreters and fully translated survey materials for a set list of 11 survey languages. These survey languages include 10 languages identified as the most frequently requested in the 2008–09 survey year plus Welsh, which must be provided for surveys conducted in Wales.¹¹ Nine of the languages had been offered in the 2008–09 study year but two new languages (Arabic and Polish) were added to the 2009–10 study.
- 4.4.9 Interviewers carried a screening card consisting of a list of the 11 languages written in English and in the language itself, as well as a short paragraph about the survey, again written in English and in each of the 11 languages. This was used to identify the language needs of selected households where English was not spoken. Letters and leaflets were available to interviewers in these survey languages (please see section 4.6.1 for further details about translated interviews).

Confidentiality

- 4.4.10 The core and screening letters assured the respondent of confidentiality by stating:

“We assure you that your answers will be treated in the ***strictest confidence***.

It will not be possible to identify any individual person from the survey findings, and the answers you give will be used for research purposes only.

No identifiable information about you or your household will be passed to government departments, local authorities or any other bodies without your consent.”

¹¹ The 2009–10 survey languages are: Arabic, Bengali, Cantonese, Gujarati, Hindi, Polish, Punjabi (Gurumukhi script), Punjabi (Urdu Script), Somali, Urdu, Welsh.

- 4.4.11 The interviewer instructions also briefed each interviewer that they could reassure the respondent about the confidentiality of their data and the research findings.

4.5 Screening procedures

- 4.5.1 In order to identify ethnic minority respondents eligible for the boost sample, the survey used two types of screening procedures: Focused enumeration screening was used in areas with medium density of ethnic minority households and direct screening was used in areas with high density of ethnic minority households (see section 2.3).

Focused enumeration screening

- 4.5.2 Interviewers with Core + focused enumeration assignments were given details of the four addresses either side of the core address on the postcode address file. They were instructed to proceed with interviewing a person aged at least 16 at the core address, and that the purpose of screening focused enumeration addresses adjacent to the core address would be to establish anyone who belongs to an ethnic minority at these addresses.
- 4.5.3 Interviewers were instructed how to screen and establish an initial screening outcome for each of the focused enumeration addresses. This could first be done by asking the respondent at the core address about the people living in the adjacent addresses using the following phraseology:

The Communities Study is interested in the views particular groups in the population have of their community and we need to take special steps to seek interviews with these people.

At [READ OUT ADDRESS BELOW], is there anyone who is of Black Caribbean, Black African, Asian, Chinese, mixed or other non-white origin? By Asian origin I mean someone whose family originally came from India, Pakistan or Bangladesh, or other Asian countries.

- 4.5.4 Interviewers were given the option to use the ethnicity showcard to help with the screening process.
- 4.5.5 If an outcome could not be established at the core address, interviewers could establish an initial screening outcome by visiting the focused enumeration addresses themselves or by visiting another adjacent property, until a definite outcome was established for each address.
- 4.5.6 Interviewers were instructed that if this – the initial screening – stage suggested there might be eligible residents living in the focused enumeration addresses,

they should visit the address in person to ask a detailed and direct screening question using a new focused enumeration contact sheet. The direct screening question was included in the focused enumeration contact sheet and interviewers were instructed to read it out exactly as it appeared.

Direct screening for the ethnic minority boost

- 4.5.7 The purpose of the ethnic minority boost was to establish eligible persons within households who were of ethnic minority origin, or, if no one of ethnic minority origin was resident, whether there was anyone resident of the Islamic faith. These addresses therefore contained two separate screening options, but were not screened as joint criteria for eligibility.
- 4.5.8 Interviewers visited each address within the ethnic minority direct screening sample and asked the detailed screening questions exactly as they appeared on the contact sheet.
- 4.5.9 The screening question for eligibility based on ethnicity was as follows:

Good morning/afternoon/evening. My name is... from Ipsos MORI/TNS-BMRB, the research company. We are carrying out a survey for the Government about the nation's households, people's attitudes towards their local area and how things can be improved. I would like to assure you that all the information will be kept in the strictest confidence by Ipsos MORI/TNS-BMRB and used for research purposes only. It will not be possible to identify any particular person, household or address in the results.

Can I just check, is there anyone aged 16 or over living in this household who belongs to any of these ethnic groups? USE SHOWCARD

May I ask which ethnic groups? USE SHOWCARD

SHOWCARD OPTIONS:

- A. Black Caribbean origin
- B. Black African origin
- C. Indian origin
- D. Pakistani origin
- E. Bangladeshi origin
- F. Chinese origin
- G. any other non-White origin
- H. Mixed origin

- 4.5.10 If respondents answered yes to any of the showcard options above, the interviewer was instructed to continue with the person selection process.
- 4.5.11 If no persons of ethnic minority origin were established at the household, the interviewer proceeded to screen on the basis of religion, in order to establish any persons in the household who were Muslim.
- 4.5.12 The screening question for eligibility based on religion was as follows:

And is there anyone aged 16 or over living in this household of the following religions, even if not currently practising? USE SHOWCARD

SHOWCARD OPTIONS:

1. No to all religions on the card
2. Christian
3. Buddhist
4. Hindu
5. Jewish
6. Sikh
7. Muslim
8. Refused to answer screening question

- 4.5.13 If faiths other than Muslim were coded, or the response was 'No to all religions on the card', interviewers coded these addresses as ineligible. If it was established that there was an eligible (age 16+) Muslim person resident, the interviewer was instructed to continue with the person selection process.

Direct screening for the Muslim boost

4.5.14 Interviewers visited each address within the Muslim sample and asked the detailed screening question exactly as it appeared on the contact sheet to establish whether anyone of Muslim faith was resident and eligible for interview.

4.5.15 The screening question was as follows:

Good morning/afternoon/evening. My name is... from Ipsos MORI/TNS-BMRB, the research company. We are carrying out a survey for the Government about the nation's households, people's attitudes towards their local area and how things can be improved. I would like to assure you that all the information will be kept in the strictest confidence by Ipsos MORI/TNS-BMRB and used for research purposes only. It will not be possible to identify any particular person, household or address in the results.

Can I just check, is there anyone aged 16 or over living in this household of the following religions, even if not currently practising? USE SHOWCARD

SHOWCARD OPTIONS:

1. No to all religions on the card
2. Christian
3. Buddhist
4. Hindu
5. Jewish
6. Sikh
7. Muslim
8. Refused to answer screening question

4.5.16 If faiths other than Muslim were coded, or the response was 'No to all religions on the card', interviewers coded these addresses as ineligible. If it was established that there was an eligible (age 16+) Muslim person resident, the interviewer was instructed to continue with the person selection process.

4.6 Screening and interviews with non-English speakers

Interpreters for screening

- 4.6.1 In order to carry out screening, interviewers were instructed to speak to an English speaker within the household where possible, with whom they could ask the screening questions. If this was not possible, interviewers were asked to record the language need using the language card as described in 4.3.1, as well as a contact telephone number. They were then able to request an interpreter on their contact sheet, who would contact the respondent, accompany the interviewer on a separate occasion.

Translated interviews

- 4.6.2 The questionnaire text, showcards and shuffle pack were translated into Welsh and the ten most commonly-requested languages in previous Citizenship Survey years. The ten languages were: Arabic, Bengali, Cantonese, Gujarati, Hindi, Polish, Punjabi (Gurmukhi script), Punjabi (Urdu script), Somali and Urdu. The translated materials were used for respondents who did not speak English, but did speak one of these survey languages.
- 4.6.3 An accredited translation company was used to translate the English-language documents, and all materials were then independently checked by a native speaker of each language to ensure consistency, accuracy and cultural equivalence.
- 4.6.4 In order to use these materials on their own, the interviewer had to speak and read the relevant language. If this was not the case, addresses were either re-issued to Citizenship Survey interviewers who did speak that language, or an interpreter was employed to accompany the original interviewer. All interpreters working on the survey received a briefing by Ipsos MORI or TNS-BMRB and were accompanied by an Ipsos MORI or TNS-BMRB interviewer.
- 4.6.5 The role of the interpreter was to read out to the respondent the translated questionnaire text from a paper document, and translate the respondent's answer into English. The interviewer then entered the answer in the (English) CAPI programme, and directed the interpreter to the next question in the script. Interpreters read from the pre-translated documents and were not allowed to re-translate the questions in any way, in order to ensure consistency amongst all respondents.
- 4.6.6 Household members were permitted to translate interview questions only if the respondent did not speak English or any of the 11 survey languages. Children aged 14 or over were allowed to translate the interview in these cases with their parent's or guardian's permission. In cases where household members were acting as translators, sensitive questions were omitted from the interview.

4.6.7 A total of 557 interviews were carried out in languages other than English, which represents three per cent of the 16,140 cases in the combined sample. Four in five translated interviews (81%) were carried out by an Ipsos MORI or TNS-BMRB interviewer or an Ipsos MORI or TNS-BMRB translator who accompanied an interviewer. Just under a fifth of translated interviews (19%) were carried out by a family member or friend. In total, 104 interviews were translated by a family member or friend. Table 4.1 summarises the number of translated interviews by language and interviewer.

Table 4.1: Translated interviews by language and interviewer

Language	Translated interviews using Ipsos MORI or TNS-BMRB interviewer	Translated interviews using an interpreter	Translated interviews using a family member or friend	Total number of translated interviews carried-out
Arabic	3	3	3	9
Bengali	49	36	7	92
Cantonese	0	1	3	4
Gujarati	26	15	4	45
Hindi	34	1	0	35
Polish	2	3	1	6
Punjabi (Gurmukhi script)	34*	0	10*	44*
Punjabi (Urdu script)	62	10	0	72
Somali	4	2	3	9
Urdu	133	5	8	146
Welsh	1	0	0	1
Other	26	3	65	94
TOTAL	374*	79	104*	557

* This includes 13 interviews carried out in Punjabi but where the script is unknown (of which, three were translated using an Ipsos MORI or TNS-BMRB interviewer, and 10 by a family member).

4.6.8 In quarter one of fieldwork, some interviews were translated in one of the 11 survey languages by a family member or friend, contravening the stipulation that household translators were to be used only for non-survey languages. In total, 29 of these interviews were conducted in quarter one. It was agreed with DCLG that these interviews could be included in the dataset. Further reminders (in the CAPI script and in a memo) were subsequently issued to ensure that this

would not happen in future quarters. In total, 65 interviews using a household translator for non-survey languages were achieved in 2009-2010.

- 4.6.9 Overall, Ipsos MORI and TNS-BMRB received 668 requests by interviewers identifying translation needs. Of these requests, 67 per cent were converted into interviews. Table 4.2 summarises the number of identified translation needs by language, and the percentage of each language that was converted into an interview.

Table 4.2 : Number of requests and completed translated interviews by language

Language	Requests for language assistance	Translated interviews completed	Percentage of requests converted to interviews
Arabic	20	9	45%
Bengali	139	92	66%
Cantonese	19	4	21%
Gujarati	57	45	79%
Hindi	36	35	97%
Polish	13	6	46%
Punjabi (Gurmukhi script)	35	31	89%
Punjabi (Urdu script)	109	72	66%
Somali	28	9	32%
Urdu	194	146	75%
Welsh	1	1	100%
TOTAL	668	450	67%

4.7 Maximising response

- 4.7.1 A number of procedures were used to maximise response rates among the households sampled for the survey.
- 4.7.2 Interviewers were instructed to make a minimum of five calls at each selected address, at different times of the day, including one evening call, and on different days of the week, including at least one call at the weekend. After quarter one, this was increased to a minimum of six calls in order to maximise response. In practice, the number of calls made to addresses was often much higher than six.

- 4.7.3 Where household contacts or selected respondents refused to give information or take part in the survey, they were asked to cite their reasons for refusal. The majority of cases where interviewers had obtained a refusal were regarded as 'soft'; that is a circumstantial rather than an absolute refusal to participate. Examples of this include contacts who said the survey would be too long or that they were too busy. These were reissued to a different interviewer who would attempt to encourage participation.
- 4.7.4 Where the interviewer was unable to make any contact at the address or dwelling unit after six or more calls (five calls in quarter one), this was recorded as a non-contact and the majority of these were reissued to another interviewer. There were other cases where final outcomes were classified as 'unproductive', for example, codes such as 'at home ill during survey period'. These were sometimes reissued to another interviewer depending on exactly what the outcome was and what the interviewer had noted about the address.
- 4.7.5 All interviewers working on the survey received news bulletins throughout the year containing tips for achieving high response, important religious dates that interviewers needed to be aware of when calling at sample point addresses and other relevant and useful information.
- 4.7.6 During quarter 3 fieldwork (1 October 2009–31 December 2009), we encouraged interviewers to begin their work earlier because the Christmas period affects fieldwork at the end of the quarter.

4.8 Interview length

- 4.8.1 The median interview length was 51 minutes. This length is calculated only on full interviews and not partial interviews.
- 4.8.2 Partial interviews are all interviews terminated before the Volunteering section of the questionnaire. In total, 53 partial interviews were completed.

Section 5

Response rates

5.1 Introduction

5.1.1 This chapter presents a discussion on response rates by sample type. Tables providing response rates for the different sample types are located at the end of this chapter.

5.2 Core sample

5.2.1 Of the 18,120 sampled core addresses, 8 per cent were classified as ineligible as they did not contain an occupied private household – five per cent were empty and one per cent comprised non-residential addresses such as businesses.

5.2.2 A total of 16,622 addresses were classified as in-scope. Of these, 27 per cent of addresses were refusals, either at the dwelling unit or by the selected person. At five per cent of addresses the interviewer was unable to make any contact, again either at the dwelling unit or with the selected person, while in six per cent of cases it was not possible to establish the eligibility of the address. At a further seven per cent of addresses, the interviewer was unable to conduct an interview because, for example, the selected person was ill or they were physically or mentally unable.

5.2.3 A total of 9,305 interviews were achieved, producing a response rate for in-scope core addresses of 56 per cent. Table 5.1 details the response rate for the core sample.

5.3 Boost sample – focused enumeration screening

5.3.1 The number of issued (i.e. enumerated) addresses was 49,760; four times the number of core addresses issued with associated focused enumeration addresses (because two addresses either side of the core address were sampled). At 96 per cent of the issued addresses, no one from an ethnic minority group was reportedly living there, the address was classified as ineligible or it was not possible to obtain the initial screening information. A total of 1,879 addresses, equivalent to four per cent of issued addresses, were initially identified for the direct screening stage.

- 5.3.2 Of the addresses then directly screened, 71 per cent contained at least one eligible adult. From these 1,337 eligible addresses, a total of 740 interviews were achieved, producing a response rate of 55%. Table 5.2 details the response rate for the ethnic minority boost sample issued for focused enumeration screening.

5.4 Boost sample – ethnic minority direct screening

- 5.4.1 A total of 30,618 addresses were issued of which nine per cent did not contain an occupied private household and were ineligible. At ten per cent of 'non deadwood' addresses, the interviewer was unable to establish whether there were any adults of ethnic minority group origin or Muslim faith resident in the dwelling unit. This was either because the people at the address or selected dwelling unit refused to answer the screening question, or because the interviewer was unable to make contact there. A further 59 per cent of cases were ineligible as they did not contain an adult of ethnic minority group origin or Muslim faith.
- 5.4.2 A total of 8,620 addresses were classified as eligible addresses. Over a fifth of eligible addresses (22%) refused to take part in the survey, whilst at 13 per cent of eligible addresses the interviewer was unable to make contact either at the selected dwelling unit or with the selected person.
- 5.4.3 A total of 4,540 interviews were achieved at eligible addresses, producing an ethnic minority response rate of 53 per cent. Table 5.3 details the response rates for the direct screening ethnic minority boost sample.

5.5 Boost sample – Muslim direct screening

- 5.5.1 A total of 19,764 addresses were issued of which eight per cent did not contain an occupied private household and were therefore ineligible. At nine per cent of 'non deadwood' addresses, the interviewer was unable to establish whether any adults of Muslim faith were resident, either because the people at the address or selected dwelling refused to answer the screening question, or because the interviewer was unable to make contact there. A further 75 per cent of cases were ineligible, as they did not contain an adult of Muslim faith.
- 5.5.2 For the 2,958 remaining eligible addresses, a fifth (20%) of addresses refused to take part in the survey, whilst at 11 per cent of eligible addresses the interviewer was unable to make contact either at the selected dwelling unit or with the selected person.

5.5.3 A total of 1,555 interviews were achieved at eligible addresses, producing a Muslim response rate of 53 per cent. Table 5.4 details the response rates for the direct screening Muslim boost sample.

Table 5.1: Response rates: core sample			
	Number	Issued cases	In-scope cases
	N	%	%
Total issued addresses	18,120	100	
Not yet built/under construction	22	0	
Demolished/derelict	61	0	
Vacant/empty housing unit	956	5	
Non-residential address (e.g. business)	179	1	
Communal establishment/institution	41	0	
Occupied but not as main residence	185	1	
Other	54	0	
Total ineligible addresses	1,498	8	
Total in-scope addresses	16,622		100
Total unknown eligibility	942		6
Refusal by phoning office	265		2
Further information about occupants refused	926		6
Refusal by selected person	2,601		16
Proxy refusal (including refusal by parents)	358		2
Broken appointment	316		2
Total refusals	4,466		27
No further contact at address/dwelling unit	467		3
Contact made at address/dwelling unit but not with 16+	13		0
No parental permission obtained – no contact with parent	3		0
No further contact with selected person	337		2
Total non contact	820		5

Table 5.1: Response rates: core sample (continued)

	Number	Issued cases	In-scope cases
	N	%	%
Unable to select person due to physical/mental ability of contact	64		0
Unable to select person due to language barrier of contact	22		0
Other reason for not selecting person	85		1
Contact made with selected person but no specific appointment made	190		1
At home ill during survey period	110		1
Away or in hospital during survey period	178		1
Physical or mentally unable/incompetent for interview	204		1
Language – inadequate English, no translator/interpreter available for interview	49		0
Other unproductive	187		1
Total other unproductive	1,089		7
Full interview	9,285		56
Partial interview	20		0
Total interviews	9,305		56

Table 5.2: Response rates: focused enumeration screening

	Number	Issued cases	Cases identified for direct screening	Cases available for direct screening	In-scope cases
	N	%	%	%	%
Issued core addresses with associated focused enumeration	12,440				
Total issued addresses	49,760	100			
Addresses identified for direct screening	1,879	4	100		
Ineligible addresses	131		7		
Ineligible (from screening)	196		10	11	
Unknown eligibility	215		11	12	
Total in-scope addresses	1,337		71	76	100
Total refusals	329		18	19	25
Total non contact	138		7	8	10
Total other unproductive	130		7	7	10
Full interview	736		39	42	55
Partial interview	4		0	0	0
Total interviews	740		39	42	55

Table 5.3: Response rates: ethnic minority direct screening

	Number	Issued cases	Non-deadwood cases	In-scope cases
	N	%	%	%
Total issued addresses	30,618	100		
Not yet built/under construction	39	0		
Demolished/derelict	152	0		
Vacant/empty housing unit	1,602	5		
Non-residential address (e.g. business)	531	2		
Communal establishment/institution	100	0		
Occupied but not as main residence	148	0		
Other	133	0		
Total ineligible addresses	2,705	9		
Total non-deadwood addresses	27,913	91		
Issued but not attempted	12		0	
Inaccessible	51		0	
Unable to locate address	151		1	
No contact made at address	1,843		7	
All information about refused	258		1	
Unable to establish eligibility due to physical/mental ability	3		0	
Unable to confirm eligibility due to language barrier	47		0	
Refusal by phoning office	22		0	
Refused to answer screening question	419		2	
Other	103		0	
Total unknown eligibility	2,909		10	
Total addresses screened	25,004		90	
No-one non-White or of Muslim faith	16384		59	
Total in-scope addresses	8,620		31	100
Further information about occupants refused	556		2	6
Refusal by selected person	862		3	10

Table 5.3: Response rates: ethnic minority direct screening (continued)				
	Number	Issued cases	Non-deadwood cases	In-scope cases
	N	%	%	%
Proxy refusal (including refusal by parents)	189		1	2
Broken appointment	296		1	3
Total refusals	1,903		7	22
No further contact at address/dwelling unit	740		3	9
Contact made at address/dwelling unit but not with 16+	17		0	0
No parental permission obtained – no contact with parent	7		0	0
No further contact with selected person	399		1	5
Total non contact	1,163		4	13
Unable to select person due to physical/mental ability of contact	18		0	0
Unable to select person due to language barrier of contact	64		0	1
Other reason for not selecting person	87		0	1
Contact made with selected person but no specific appointment	180		1	2
At home ill during survey period	38		0	0
Away or in hospital during survey period	139		0	2
Physical or mentally unable/incompetent for interview	55		0	1
Language – no translator/interpreter available for interview	182		1	2
Other unproductive	251		1	3
Total other unproductive	1,014		4	12
Full interview	4,519		16	52
Partial interview	21		0	0
Total interviews	4,540		16	53

Table 5.4 Response rates: Muslim direct screening

	Number	Issued cases	Non-deadwood cases	In-scope cases
	N	%	%	%
Total issued addresses	19,764	100		
Not yet built/under construction	35	0		
Demolished/derelict	116	1		
Vacant/empty housing unit	909	5		
Non-residential address (e.g. business)	336	2		
Communal establishment/institution	26	0		
Occupied but not as main residence	20	0		
Other	115	1		
Total ineligible addresses	1,557	8		
Total non-deadwood addresses	18,207	92	100	
Issued but not attempted	35		0	
Inaccessible	130		1	
Unable to locate address	51		0	
No contact made at address	974		5	
All information about refused	64		0	
Unable to establish eligibility due to physical/mental ability	37		0	
Unable to confirm eligibility due to language barrier	40		0	
Refusal by phoning office	6		0	
Refused to answer screening question	231		1	
Other	10		0	
Total unknown eligibility	1,578		9	
Total addresses screened	16,629		91	
No-one of Muslim faith	13,671		75	
Total in-scope addresses	2,958		16	100
Further information about occupants refused	172		1	6
Refusal by selected person	252		1	9
Proxy refusal (including refusal by parents)	90		1	3

Table 5.4 Response rates: Muslim direct screening (continued)				
	Number	Issued cases	Non-deadwood cases	In-scope cases
	N	%	%	%
Broken appointment	91		0	3
Total refusals	605		3	20
No further contact at address/dwelling unit	196		1	7
Contact made at address/dwelling unit but not with 16+	4		0	0
No parental permission obtained – no contact with parent	0		0	0
No further contact with selected person	123		1	4
Total non contact	323		2	11
Unable to select person due to physical/mental ability of contact	6		0	0
Unable to select person due to language barrier of contact	45		0	2
Other reason for not selecting person	34		0	1
Contact made with selected person but no specific appointment	66		0	2
At home ill during survey period	18		0	1
Away or in hospital during survey period	49		0	2
Physical or mentally unable/incompetent for interview	20		0	1
Language – no translator/interpreter available for interview	134		1	5
Other unproductive	103		0	3
Total other unproductive	475		2	16
Full interview	1,543		9	52
Partial interview	12		0	0
Total interviews	1,555		9	53

Section 6

Data Processing

6.1 Editing

- 6.1.1 Where clearly relevant, the CAPI program specified numerical ranges to ensure answers were sensible. Logic checks were also scripted to check answers that were not feasible or were incorrect, e.g. if an interviewer coded that a respondent was not Muslim when conducting an interview in the Muslim boost sample.

6.2 Coding

- 6.2.1 Post-interview coding was undertaken by members of the Ipsos MORI and TNS-BMRB coding departments, using identical codeframes. The codeframes drew upon previous survey years' codeframes as well as an agreed codeframe for new questions. The coding departments coded verbatim responses recorded at open and 'other – specify' questions, and occupation and socio-economic class.

Open and 'other – specify' questions

- 6.2.2 Researchers, in partnership with DCLG, developed a codeframe to categorise verbatim responses to the three open questions: EDMixpr (introduced in the 2007–08 survey), Verelol and Vthi (both introduced in 2009–10 survey). EDMixpr asked all respondents who felt that their local area was not cohesive (i.e. disagreed that people from different backgrounds got on well together in their local area) what sort of things prevent people from different backgrounds getting on well together. Verelol asked in what circumstances respondents felt it would be right to use violent extremism in the name of religion to protest or achieve a goal. Vthi asked those who had seen and read materials encouraging violent extremism what they thought of these materials.
- 6.2.3 In addition, researchers extended the code frames (where necessary) of 'other – specify' questions, where the responses that did not fit into a pre-code were above 10 per cent of the total number of responses.

Occupation and socio-economic class

6.2.4 Occupation details were collected for the respondent and the household reference person where this was not the respondent. Occupations were coded according to the Standard Occupational Classification (SOC2000). This was carried out by coders at Ipsos MORI and TNS-BMRB using the computer-assisted coding process CASCOT.

6.3 Derived variables

6.3.1 A list of the main derived variables are provided in Annex F (under separate cover to this report).

6.3.2 The following geo-demographic variables were added to the data:

- Government Office Region;
- Local authority;
- ACORN¹² classification;
- Urban/rural indicator;
- Percentage of households in the ward headed by someone from a non-white ethnic minority group;
- Index of Multiple Deprivation for England (2007);
- Index of Multiple Deprivation for Wales (2007);
- Office for National Statistics classification of local authorities;
- Office for National Statistics classification of health authorities; and
- Police service area.

6.3.3 Some geo-demographic variables are not included in the publicly available dataset (see paragraph 6.4.4). This is to prevent the possibility of survey respondents being identified through this detailed geographic information.

6.4 Data outputs

6.4.1 DCLG received a full cumulative SPSS dataset including derived variables and additional variables each quarter.

6.4.2 DCLG also received a set of Excel tables providing cumulative data on a subset of the dataset for each quarter.

¹² ACORN is a geodemographic classification of the UK population, ACORN codes are allocated to postcodes and describe the predominate characteristics of the population within that postcode. More information can be found on the following website <http://www.caci.co.uk/acorn/whatis.asp>

- 6.4.3 The data are also publicly available via the UK ESRC Data Archive in SPSS format: www.data-archive.ac.uk .
- 6.4.4 The publicly available dataset does not include any detailed geo-demographic variables i.e., ward codes, local authority codes, urban codes, health area name and subgroups, police service area codes and descriptions, housing association groups and subgroups and ethnic proportion of ward, as well as other administrative variables.
- 6.4.5 A guide to using the public SPSS dataset is provided in Chapter 9.

Section 7

Weighting

7.1 Introduction

7.1.1 The following four weights have been calculated for the 2009–10 survey data:¹³

- A household weight for the core sample;
- An individual weight for the core sample;
- A household weight for the combined core, ethnic minority and Muslim boost samples; and
- An individual weight for the combined core, ethnic minority and Muslim boost samples.

7.1.2 Both the core and combined sample weights adjust the sample for differences in response rates and take account of dwelling unit and individual selection probabilities. The combined sample weights also adjust the sample for differences in screening, and for unequal address selection probabilities (because of oversampling of addresses in ethnic minority and Muslim boost areas).

7.1.3 In a change¹⁴ from the 2008–09 study, the non-response weights for the combined sample were calculated separately for each of its constituent sample types (core, ethnic minority boost excluding focused enumeration, and Muslim), before the design and final calibration¹⁵ weighting was applied. Moreover the non-response weights for the ethnic minority and Muslim boosts were separated into screening non-response and cooperation non-response, each of which was calculated separately. Non-response weights were not calculated for the focused enumeration element of the ethnic minority boost.¹⁶ In another change from 2008–09, the weights were calculated for each quarter and then combined across quarters rather than being calculated based on one year's worth of data as they were previously.

¹³ The weighting procedure is based on that used by the Office of National Statistics on the 2003 survey (Green and Farmer, 2004).

¹⁴ In the 2008–09 Citizenship Survey, the non-response modelling for the combined weight was calculated based on the whole sample rather than separately for each sample type.

¹⁵ See paragraphs 7.2.9 and 7.2.10 for explanation of calibration weighting.

¹⁶ Focused enumeration is not conducive to non-response weighting due to the nature by which eligibility is established by visiting addresses adjacent to the core address.

7.1.4 The core weight should be used for any estimates using core sample data relating to the general population or sub-groups unrelated to ethnic minority or Muslim groups, whereas the combined weight should be used for any analysis using the combined sample and relating to estimates for ethnic minority or Muslim groups or sub-groups relating to ethnic minority or Muslim groups. An individual and household weight has been generated for each sample; the core and combined. The recommended application of annual weights is summarised in Table 7.1.

Weight	Type of estimate	Base
WtCHH (core sample household weight)	Household estimates for whole sample and for sub-groups apart from (i) ethnic minority/Muslim groups and (ii) sub-groups relating to ethnic minority/Muslim groups e.g. Country of Birth	Core sample only (unweighted base = 9,305)
WtCInds (core sample individual weight)	Individual estimates for whole sample and for sub-groups apart from (i) ethnic minority/Muslim groups and (ii) sub-groups relating to ethnic minority/Muslim groups e.g. Country of Birth	Core sample only (unweighted base = 9,305)
WtCombHH (combined sample household weight)	Household estimates for ethnic minority/Muslim groups and sub-groups relating to ethnic minority and/or Muslim groups e.g. Country of Birth	Combined sample only (unweighted base = 16,140)
WtFInds (combined sample Individual weight)	Individual estimates for ethnic minority/Muslim groups and sub-groups relating to ethnic minority and/or Muslim groups e.g. Country of Birth	Combined sample only (unweighted base = 16,140)

Quarterly weights

- 7.1.5 At the end of each quarter the data were weighted to enable quarterly estimates to be produced. The weighting method outlined below was used to generate core and combined weights for each quarter. Unlike the 2008–09 study, the quarterly weights were used in the generation of weights for the yearly data.¹⁷ To avoid over-representing any quarter in the final yearly core and combined datasets, an adjustment was made to the quarterly weights so that the sum of the weights for each quarter were equivalent, thereby ensuring no single quarter had more influence over the survey estimates than any other (if for example there was a larger sample size in one quarter).
- 7.1.6 As the process of deriving the weights is the same each quarter (the only difference being the outputs from the modelling), we have only provided non-response models based on the fourth quarter of 2009–10. These are reproduced in sections 7.2 and 7.3 to illustrate the quarterly weighting process.

7.2 Calculation of core sample weights

- 7.2.1 The core sample and weights should be used for generating household and individual estimates for the general population, including estimates for whole sample or subgroups apart from (i) ethnic minority or Muslim groups and (ii) sub-groups relating to ethnic minority or Muslim groups.
- 7.2.2 There are two sets of weights for the core sample for analyses at (a) the household, and (b) the individual level.

Calculation of household weight for the core sample

- 7.2.3 This comprises the dwelling unit selection weight (w_1) and the household non-response rate (w_2), the product of which is calibrated to produce the final household weight for the core sample.

Dwelling unit selection weight

- 7.2.4 At each contacted address the interviewer established the number of dwelling units. Whilst most addresses contained a single dwelling unit, at a small proportion of addresses (<2%) there were multiple dwelling units. In such cases the interviewer used the Kish grid¹⁸ to select a single dwelling unit for inclusion in the survey. The dwelling unit selection weight (w_1) adjusts for this selection and is equivalent to the number of dwelling units at the selected address. This weight has been trimmed¹⁹ to a maximum of four to avoid any large values.

¹⁷ In 2008-9 quarterly weights were produced for analysis on individual quarters but were not used in the generation of weights for the yearly data; the final weights for the yearly data were generated from scratch.

¹⁸ A computer-generated Kish grid was provided on each contact sheet.

¹⁹ Trimming is the process by which we set a maximum threshold on the number of dwelling units at an address and the number of eligible people in a dwelling unit. It is used to reduce the adverse effects of having respondents in the survey with large weights.

Household non-response weight

- 7.2.5 The household non-response weight is based on a logistic regression model of the response behaviour of households in the core sample. The regression was run on unweighted data due to the equal probability of selection sample design. This model generates the probability of a household participating in the survey given their characteristics (based on a set of predictor variables described in paragraph 7.2.6 below). The household non-response weight (w_2) is then calculated as the inverse of the predicted probabilities. Hence households that were of a type that were more reluctant to take part will have a smaller predicted probability and a larger weight.
- 7.2.6 In order to ensure consistency across different years, the predictor variables used for 2009-10 were the same as those used previously. These variables have been shown to best describe variations in likelihood of response.²⁰ The predictor variables used in the model were thus: Government Office Region, ACORN group (16 categories) and quintiles of the proportion of the Output Area's population belonging to a non-white ethnic minority group, based on data from the 2001 Census.²¹
- 7.2.7 Across the four models generated for each quarter of the survey, response was consistently lower (although not always significantly lower) in London relative to the other Government Office Regions, in areas with high density of ethnic minority groups relative to areas with low density of ethnic minority groups and in ACORN group E (Educated Urbanites)
- 7.2.8 The full model for quarter 4 is presented in Table 7.2. The coefficients in the table relate to how much the predicted probability of response increases (or decreases, if the coefficient is negative) relative to the reference category when a household has that particular characteristic. The expected probability of response can be generated for a household by using these coefficients with the corresponding values of the predictor variables for that particular household. The household non-response weights were trimmed to remove a small number of outliers.²² Large weights relative to very small weights increase the variance in the survey estimates and so it is common practice in weighting survey data to accept a small amount of bias by trimming the weights for increased precision.

²⁰ The 2003 Citizenship survey Technical Report Chapter 7.2 (<http://www.esds.ac.uk/doc/5087/mrdoc/pdf/5087userguide.pdf>)

²¹ It should be noted that some bias/error may be introduced where changes to the population have occurred since 2001.

²² Outliers are identified as those being in the top percentile of the non-response weight distribution.

Table 7.2: Quarter 4 core sample household non-response model

	B	S.E.	Wald	df	Sig.	Exp(B)
Government Office Region			17.06	9	0.048	
East Midlands					(baseline)	
East of England	0.00	0.15	0.00	1	0.986	1.00
London	-0.27	0.17	2.55	1	0.110	0.77
North East	0.14	0.18	0.56	1	0.455	1.15
North West	0.13	0.14	0.88	1	0.348	1.14
South East	0.19	0.14	1.85	1	0.173	1.21
South West	0.20	0.15	1.73	1	0.188	1.22
Wales	0.42	0.18	5.32	1	0.021	1.52
West Midlands	0.10	0.15	0.48	1	0.489	1.11
Yorkshire and The Humber	0.07	0.15	0.19	1	0.659	1.07
Acorn Group			38.78	17	0.002	
A Wealthy Executives					(baseline)	
B Affluent Greys	-0.35	0.15	5.15	1	0.023	0.71
C Flourishing Families	-0.16	0.15	1.04	1	0.307	0.86
D Prosperous Professionals	-0.17	0.26	0.43	1	0.510	0.84
E Educated Urbanites	-0.82	0.20	17.51	1	0.000	0.44
F Aspiring Singles	-0.40	0.21	3.74	1	0.053	0.67
G Starting Out	-0.45	0.20	5.21	1	0.022	0.64
H Secure Families	-0.35	0.13	6.65	1	0.010	0.71
I Settled Suburbia	-0.32	0.16	4.33	1	0.037	0.72
J Prudent Pensioners	-0.38	0.18	4.44	1	0.035	0.68
K Asian Communities	0.05	0.35	0.02	1	0.897	1.05
L Post-Industrial Families	-0.08	0.19	0.18	1	0.669	0.92
M Blue-collar Roots	-0.02	0.16	0.01	1	0.910	0.98
N Struggling Families	-0.09	0.14	0.41	1	0.521	0.91
O Burdened Singles	0.17	0.19	0.83	1	0.361	1.19
P High-Rise Hardship	-0.30	0.25	1.42	1	0.234	0.74
Q Inner City Adversity	-0.44	0.26	2.87	1	0.090	0.64
Unclassified/missing	-0.07	0.69	0.01	1	0.915	0.93

Table 7.2: Quarter 4 core sample household non-response model (continued)

	B	S.E.	Wald	df	Sig.	Exp(B)
% non-white ethnic minority population by Output Area (Census 2001)			8.50	4	0.075	
Bottom Quintile					(baseline)	
Second Quintile	0.05	0.10	0.22	1	0.639	1.05
Third Quintile	-0.09	0.10	0.75	1	0.386	0.92
Fourth Quintile	-0.18	0.11	2.86	1	0.091	0.83
Top Quintile	-0.29	0.13	4.69	1	0.030	0.75
Constant	0.51	0.17	9.30	1	0.002	1.66

Please see notes below.²³

Final calibrated household weight

7.2.9 The household weight is calculated as the product of the dwelling unit selection weight and the household non-response weight ($w_1 \times w_2$). The final step was to calibrate²⁴ this household weight.

7.2.10 Calibration weighting is a technique that creates weights which, when applied to survey data, give survey estimates that match the population estimates for certain key variables. It corrects for any differences due to random chance in the selection process and the uncorrected effect of differential non-response between the (weighted) achieved sample and the population profile.

7.2.11 Calibration weighting allows household weights to be generated that are based on the characteristics of the household members. This means households can be weighted using external information about individuals, which is more reliable and readily available than external information about households. The information used was the 2008 Office of National Statistics mid-year household population estimates for England and Wales.²⁵ The method means

²³ Notes:

1. The response is 1 = household response, 0 = no household response.
2. The model $R^2 = 0.033$ (Cox and Snells).
3. B is the estimate coefficient with standard error S.E.
4. The Wald-test measures the impact of the categorical variable on the model with the appropriate number of degrees of freedom df. If the test is significant ($\text{sig} < 0.05$) then the categorical variable is considered to be 'significantly associated' with the response variable.
5. The Wald test for each level of the categorical variable is also shown. This tests the difference between that level and the baseline category.

²⁴ The calibration was carried out in g-Calib, a macro program run in SPSS which adjusts the margins of a contingency table of survey estimates to match the known population margins. See Deville J-C & Sarndal C-E (1992)

²⁵ The 2008 household population estimates are experimental statistics. They are not National Statistics as they do not meet the stringent requirements made of National Statistics data. Whilst these estimates are a better representation of the population covered by our sample, their experimental nature may mean there are issues of accuracy or quality. The estimates have been used in a very aggregated form as weighting totals; less aggregated totals would be less reliable. When the Quarter 1 weights were calculated the 2008 population estimates were the latest estimates available. The subsequent quarters and combined quarter weights used the same totals to make them comparable.

the calibration weight for a particular household depends upon the age/sex profiles of the household members, which reflects the relationship between the likelihood of household members (and hence the household) to participate and their age and sex. Including region ensured the calibration weights also took account of the differential response by region identified in the household non-response model.

- 7.2.12 The population estimates used for the calibration were age/sex (16 categories) and Government Office Region (10 categories, including Wales) – see Tables 7.3 and 7.4. The final household weight for the core sample across all 4 quarters (WtCHH) is summarised in Table 7.5.
- 7.2.13 This weight (WtCHH) should be applied in the software package when interrogating the dataset for any household level estimates e.g. proportion of single parent households in England and Wales.

Table 7.3 : 2008 mid-year household population estimates by age and sex

Age by sex	N	%
Male 0-15	5,234,500	9.6
Male 16-24	3,387,800	6.2
Male 25-34	3,550,200	6.5
Male 35-44	4,038,900	7.4
Male 45-54	3,552,900	6.5
Male 55-64	3,162,300	5.8
Male 65-74	2,170,400	4.0
Male 75+	1,683,300	3.1
Female 0-15	4,988,200	9.2
Female 16-24	3,212,000	5.9
Female 25-34	2,480,900	6.4
Female 35-44	4,089,500	7.5
Female 45-54	3,633,500	6.7
Female 55-64	3,284,500	6.0
Female 65-74	2,385,800	4.4
Female 75+	2,585,100	4.7
All	54,439,800	100

Table 7.4: 2008 mid-year household population estimates by Government Office Region

Government Office Region	N	%
North East	2,575,500	4.7
North West	6,875,700	12.6
Yorkshire and The Humber	5,213,200	9.6
East Midlands	4,433,000	8.1
West Midlands	5,411,100	9.9
East	5,728,700	10.5
London	7,619,800	14.0
South East	8,380,100	15.4
South West	5,209,200	9.6
Wales	2,993,400	5.5
All	54,439,700²⁶	100

Table 7.5: Quarter 1 to quarter 4 summary of final household weight (core sample)

WtCHH	N	Range	Minimum and maximum	Mean	Median	5th and 95th percentile
Household Characteristics						
North East	537	8.1	0.18 – 8.32	1.02	0.90	0.58 – 1.68
North West	1,284	7.7	0.03 – 7.71	0.89	0.85	0.53 – 1.62
Yorkshire and The Humber	907	11.5	0.03 – 11.56	1.10	1.05	0.69 – 1.60
East Midlands	788	4.4	0.03 – 4.41	0.88	0.90	0.32 – 1.36
West Midlands	933	5.7	0.03 – 5.70	1.08	1.07	0.65 – 1.53
East	934	7.2	0.19 – 7.40	1.16	1.07	0.70 – 1.80
London	975	6.3	0.05 – 6.37	1.23	1.22	0.13 – 2.30
South East	1,407	6.5	0.00 – 6.49	0.81	0.84	0.15 – 1.33
South West	947	3.6	0.04 – 3.61	0.91	0.88	0.60 – 1.32
Wales	593	3.9	0.03 – 3.98	1.09	0.95	0.59 – 1.85
All	9,305	11.6	0.00 – 11.56	1.00	0.94	0.38 – 1.70

²⁶ The 2008 mid-year population estimates for age by gender and Government Office Region were different by a total of 100 people. Therefore, when weighting the data, minor adjustments proportional to the population size of each Government Office Region were made, so that the total population counts by age and gender and Government Office Region were equal to 54,439,800.

Calculation of individual weight for core sample

7.2.14 This comprises the dwelling unit selection weight (w_1 , as above), the household non-response rate (w_2 , as above), and the individual selection weight (w_3 , see below), the product of which is then calibrated to produce the final individual weight for the core sample.

Individual selection weight

7.2.15 At each selected dwelling unit one individual was selected at random from all the adults in the household aged 16 or over. The individual selection weights (w_3) are generated based on the number of eligible individuals in the household. Without these weights, individuals in larger households would be under-represented in the sample.

7.2.16 To avoid excessively large weights having an undue influence on the estimates the individual selection weight was trimmed to a maximum of four for the core sample.

Final calibrated individual weight

7.2.17 The individual weight for the core sample is the product of the dwelling unit selection, household non-response and individual selection weights ($w_1 \times w_2 \times w_3$). This weight was then calibrated. Unlike the household calibration weighting, which used information for all household members, in this case only information about the selected individual was used. Hence, the characteristics of the (weighted) achieved sample of individuals were adjusted to match the population of England and Wales aged 16 and over, according to the 2008 mid-year household population estimates, as individuals were only eligible for the survey if they were aged 16 or over.

7.2.18 The population estimates used for the calibration were age/sex (14 categories) and Government Office Region (10 categories, including Wales). The population figures are given in Tables 7.6 and 7.7. The final individual weight for the core sample for all 4 quarters ($WtCInds$) is summarised in Table 7.8.

7.2.19 This weight ($WtCInds$) should be applied in the software package when interrogating the dataset for any individual level estimates e.g. proportion of adults 16 years and over in England and Wales with one or more children.

Table 7.6: 2008 mid-year household population estimates by age and sex		
Age by sex	N	%
Male 16-24	3,387,800	7.7
Male 25-34	3,550,200	8.0
Male 35-44	4,038,900	9.1
Male 45-54	3,552,900	8.0
Male 55-64	3,162,300	7.2
Male 65-74	2,170,400	4.9
Male 75+	1,683,300	3.8
Female 16-24	3,212,000	7.3
Female 25-34	2,480,900	7.9
Female 35-44	4,089,500	9.2
Female 45-54	3,633,500	8.2
Female 55-64	3,284,500	7.4
Female 65-74	2,385,800	5.4
Female 75+	2,585,100	5.8
All	44,217,100	100

Table 7.7: 2008 mid-year household population estimates by Government Office Region		
Government Office Region	N	%
North East	2,114,200	4.8
North West	5,575,600	12.6
Yorkshire and The Humber	4,242,500	9.6
East Midlands	3,616,800	8.2
West Midlands	4,359,800	9.9
East	4,642,300	10.5
London	6,147,800	13.9
South East	6,790,700	15.4
South West	4,287,000	9.7
Wales	2,440,300	5.5
All	44,217,000²⁷	100

²⁷ The 2008 mid-year population estimates for age by gender and Government Office Region were different by a total of 100 people. Therefore when weighting the data, minor adjustments proportional to the population size of each Government Office Region were made, so that the total population counts by age and gender and Government Office Region were equal to 44,217,100.

Table 7.8: Quarter 1 to quarter 4 summary of final individual weight (core sample)

WtCInds	N	Range	Minimum and maximum	Mean	Median	5th and 95th percentile
Individual Characteristics						
Government Office Region						
North East	537	4.59	0.33 – 4.92	0.83	0.79	0.37 – 1.62
North West	1,284	3.23	0.35 – 3.58	0.91	0.86	0.40 – 1.87
Yorkshire and The Humber	907	5.11	0.36 – 5.47	0.98	0.91	0.43 – 1.99
East Midlands	788	5.66	0.33 – 5.99	0.97	0.91	0.43 – 1.88
West Midlands	933	3.01	0.34 – 3.35	0.98	0.91	0.44 – 1.98
East	934	2.99	0.34 – 3.33	1.04	0.97	0.45 – 2.04
London	975	9.37	0.37 – 9.73	1.33	1.14	0.52 – 2.82
South East	1,407	4.88	0.37 – 5.25	1.02	0.94	0.44 – 2.06
South West	947	6.35	0.33 – 6.68	0.95	0.89	0.41 – 1.96
Wales	593	2.57	0.32 – 2.89	0.87	0.80	0.38 – 1.75
Sex						
Male	4,173	9.40	0.33 – 9.73	1.09	0.97	0.42 – 2.25
Female	5,132	4.93	0.32 – 5.25	0.93	0.86	0.42 – 1.89
Age group						
16-24	797	9.31	0.42 – 9.73	1.74	1.63	0.61 – 3.01
25-34	1,313	4.79	0.39 – 5.19	1.13	1.06	0.46 – 2.23
35-44	1,718	4.91	0.34 – 5.25	1.00	0.95	0.43 – 1.85
45-54	1,479	3.68	0.32 – 4.00	1.02	0.97	0.44 – 1.93
55-64	1,570	5.66	0.33 – 5.99	0.86	0.86	0.40 – 1.58
65-74	1,266	4.26	0.33 – 4.59	0.76	0.79	0.39 – 1.24
75+	1,162	2.02	0.36 – 2.38	0.77	0.64	0.42 – 1.34
All	9,305	11.6	0.00 – 11.56	1.00	0.94	0.38 – 1.70

Quarterly adjustment

7.2.20 Unlike the 2008–09 study, the quarterly weights were used in the generation of weights for the annual data. To avoid over-representing any quarter when calculating estimates from the survey based on a whole year of data a quarterly adjustment was made. The quarterly adjustment ensured that the sum of the

weights for each of the four quarters were equal, e.g. the sample size for the core respondents across the year is 9,305, therefore the individual weights for respondents in each quarter summed to $9,305/4 = 2,326.25$.

7.3 Calculation of combined sample weights

- 7.3.1 The combined sample and weights should be used for any analyses of households or individuals by ethnic minority group and/or Muslim groups or sub-groups relating to ethnicity and religion, e.g. country of birth.
- 7.3.2 Two sets of weights, household and individual-level, were required for analysis of the combined core and boost samples.

Calculation of household weight for combined sample

- 7.3.3 This comprises the household screening non-response weight, the address selection weight, the household (cooperation) non-response weight and the dwelling unit selection weight, the product of which is calibrated to produce the final household weight for the combined sample.
- 7.3.4 In a change to the 2008–09 study, a Muslim sample boost was included in 2009–10 in addition to the ethnic minority boost (similar to that of 2008-09) (for more information please see Chapter 2).
- 7.3.5 In another change to the 2008-09 study, initial weightings (prior to calibration) of the ethnic minority and Muslim boost samples were done separately from one another and from the other constituent of the combined sample – the core sample. The processes involved for the two boost samples are very similar and the section below describes the general process for both, identifying where they differ. Where model output is presented this comes from the Quarter 4 ethnic minority boost.

Household screening non-response weight

- 7.3.6 For the ethnic minority and Muslim boosts, screening non-response and cooperation non-response were modelled separately. All modelling was done at the household level. The household screening non-response weights were based on a logistic regression model run on unweighted data for the ethnic minority boost and weighted data for the Muslim boost.²⁸ The model was used to generate a predicted probability of being screened for each household given their characteristics (based on a set of predictor variables described in para 7.2.5 above). The predictor variables used in the model were Government Office

²⁸ The screening non-response model was run on weighted data for the Muslim Boost to account for different probabilities of selection dependent on whether you were a Muslim living in Output Areas with between 2.5 and under 10 per cent Muslim population or a Muslim living in Output Areas with 10 per cent or more of the population Muslim (for more information on why we weighted the data before modeling see Robert M. Groves and Mick P. Couper, *Non-response in Household Interview Surveys, chapter 11 section 9*).

Region, ACORN group (16 categories) and quintiles of the proportion of the output area's population belonging to a non-white ethnic minority group, based on data from the 2001 Census. The household screening non-response weight (w1) was generated as the inverse of the selected probabilities from the logistic regression model. The full model for the ethnic minority boost in Quarter 4 is presented in Table 7.9. The household screening non-response weights were trimmed to remove a small number of outliers.

- 7.3.7 The model shows those households more likely to be screened reside in the South East, West Midlands or the East of England. They reside in a postcode categorised as ACORN group A or B (Wealthy Achievers/Affluent Greys) and in local areas where there is a lower than average proportion of non-white residents.

Table 7.9: Quarter 4 ethnic minority sample household screening non-response model

	B	S.E.	Wald	df	Sig.	Exp(B)
Government Office Region			34.81	7	0.000	
East Midlands					(baseline)	
East of England	0.94	0.42	5.04	1	0.025	2.55
London	0.19	0.21	0.82	1	0.365	1.21
North West	-0.01	0.26	0.00	1	0.968	0.99
South East	1.08	0.32	11.72	1	0.001	2.94
Wales	0.26	0.38	0.45	1	0.501	1.30
West Midlands	0.97	0.27	12.50	1	0.000	2.63
Yorkshire and The Humber	0.08	0.26	0.09	1	0.758	1.08
Acorn Group			70.95	16	0.000	
A/B. Wealthy Executives/ Affluent Greys					(baseline)	
C. Flourishing Families	-0.43	0.93	0.21	1	0.645	0.65
D. Prosperous Professionals	-1.48	0.78	3.62	1	0.057	0.23
E. Educated Urbanites	-1.74	0.73	5.67	1	0.017	0.18
F. Aspiring Singles	-1.14	0.73	2.42	1	0.120	0.32
G. Starting Out	-1.41	0.77	3.34	1	0.068	0.24
H. Secure Families	-1.01	0.74	1.86	1	0.172	0.37
I. Settled Suburbia	-0.33	1.02	0.10	1	0.747	0.72

Table 7.9: Quarter 4 ethnic minority sample household screening non-response model

	B	S.E.	Wald	df	Sig.	Exp(B)
J. Prudent Pensioners	-1.80	0.78	5.33	1	0.021	0.17
K. Asian Communities	-1.11	0.74	2.22	1	0.136	0.33
L. Post-Industrial Families	-1.53	0.75	4.18	1	0.041	0.22
M. Blue-collar Roots	-0.19	0.82	0.06	1	0.814	0.82
N. Struggling Families	-0.36	0.76	0.23	1	0.634	0.70
O. Burdened Singles	-1.14	0.77	2.23	1	0.136	0.32
P. High-Rise Hardship	-2.10	0.76	7.68	1	0.006	0.12
Q. Inner City Adversity	-1.30	0.73	3.13	1	0.077	0.27
Unclassified/missing	-1.63	1.05	2.40	1	0.121	0.20
% non-white ethnic minority population by Output Area (Census 2001)			25.04	4	0.000	
Bottom Quintile					(baseline)	
Second Quintile	-0.40	0.15	6.82	1	0.009	0.67
Third Quintile	-0.61	0.15	15.95	1	0.000	0.54
Fourth Quintile	-0.71	0.16	19.77	1	0.000	0.49
Top Quintile	-0.79	0.18	20.09	1	0.000	0.45
Constant	3.78	0.74	25.88	1	0.000	43.95

Notes:

1. The response is 1 = screener achieved at household, 0 = no screener achieved.
2. The model $R^2 = 0.029$ (Cox and Snells).
3. B is the estimate coefficient with standard error S.E.
4. The Wald-test measures the impact of the categorical variable on the model with the appropriate number of degrees of freedom df. If the test is significant ($\text{sig} < 0.05$) then the categorical variable is considered to be 'significantly associated' with the response variable.
5. The Wald test for each level of the categorical variable is also shown. This tests the difference between that level and the baseline category.

Address selection weight

7.3.8 An address selection weight (w_2) was generated to combat the effects of over sampling of addresses in areas used for the ethnic minority and Muslim boosts. The address selection weight is conditional on eligibility and varies according to the route the address takes into the sample; which sample the household was selected through and which samples it was eligible for. For example, an eligible Muslim of Asian ethnicity residing in a ward with 18 per cent+ of the population black and minority ethnic and in an output area within that ward with 10 per cent+ Muslim population could have come into the sample via either the core,

ethnic minority or Muslim boost samples, whilst an eligible Asian non-Muslim living in the same area could have only come into the sample via the core or ethnic minority boost samples. Whether or not an address contained at least one member from an ethnic minority or Muslim group must be known for the address selection weight to be generated, so it can only be calculated after the address has been contacted.

7.3.9 The address selection weights for Quarter 4 are presented in Table 7.10. Please note these weights are the same regardless of which quarter of the year we are looking at.

Table 7.10: Quarter 4 probability of address being selected for the combined sample

Stratum	Occupants characteristics	Probability of address selection	Address sampling weight
18%+ non-white/10%+ Muslim	non-white Muslim	0.00484	207
18%+ non-white/10%+ Muslim	non-white non-Muslim	0.00247	404
18%+ non-white/10%+ Muslim	White Muslim	0.00484	207
18%+ non-white/10%+ Muslim	White non-Muslim	0.00019	5,203
18%+ non-white/2.5 to <10% Muslim	non-white Muslim	0.00304	329
18%+ non-white/2.5 to <10% Muslim	non-white non-Muslim	0.00247	404
18%+ non-white/2.5 to <10% Muslim	White Muslim	0.00304	329
18%+ non-white/2.5 to <10% Muslim	White non-Muslim	0.00019	5,203
18%+ non-white/< 2.5% Muslim	non-white Muslim	0.00247	404
18%+ non-white/< 2.5% Muslim	non-white non-Muslim	0.00247	404
18%+ non-white/< 2.5% Muslim	White Muslim	0.00247	404
18%+ non-white/< 2.5% Muslim	White non-Muslim	0.00019	5,203
1 to < 18% non-white/10%+ Muslim	non-white Muslim	0.00333	301
1 to < 18% non-white/10%+ Muslim	non-white non-Muslim	0.00096	1,042
1 to < 18% non-white/10%+ Muslim	White Muslim	0.00256	391
1 to < 18% non-white/10%+ Muslim	White non-Muslim	0.00019	5,203
1 to < 18% non-white/2.5 to <10% Muslim	non-white Muslim	0.00152	657

Table 7.10: Quarter 4 probability of address being selected for the combined sample (continued)

Stratum	Occupants characteristics	Probability of address selection	Address sampling weight
1 to < 18% non-white/2.5 to <10% Muslim	non-white non-Muslim	0.00096	1,042
1 to < 18% non-white/2.5 to <10% Muslim	White Muslim	0.00076	1,324
1 to < 18% non-white/2.5 to <10% Muslim	White non-Muslim	0.00019	5,203
1 to < 18% non-white/< 2.5% Muslim	non-white Muslim	0.00096	1,042
1 to < 18% non-white/< 2.5% Muslim	non-white non-Muslim	0.00096	1,042
1 to < 18% non-white/< 2.5% Muslim	White Muslim	0.00019	5,203
1 to < 18% non-white/< 2.5% Muslim	White non-Muslim	0.00019	5,203
< 1% non-white/10%+ Muslim	non-white Muslim	0.00256	391
< 1% non-white/10%+ Muslim	non-white non-Muslim	0.00019	5,203
< 1% non-white/10%+ Muslim	White Muslim	0.00256	391
< 1% non-white/10%+ Muslim	White non-Muslim	0.00019	5,203
< 1% non-white/2.5 to <10% Muslim	non-white Muslim	0.00076	1,324
< 1% non-white/2.5 to <10% Muslim	non-white non-Muslim	0.00019	5,203
< 1% non-white/2.5 to <10% Muslim	White Muslim	0.00076	1,324
< 1% non-white/2.5 to <10% Muslim	White non-Muslim	0.00019	5,203
< 1% non-white/< 2.5% Muslim	non-white Muslim	0.00019	5,203
< 1% non-white/< 2.5% Muslim	non-white non-Muslim	0.00019	5,203
< 1% non-white/< 2.5% Muslim	White Muslim	0.00019	5,203
< 1% non-white/< 2.5% Muslim	White non-Muslim	0.00019	5,203

Household (cooperation) non-response weight

7.3.10 The next step was to model refusal behaviour of eligible, screened in, households to get the household (cooperation) non-response weight. Refusals were modelled using logistic regression and run on weighted data. For the ethnic minority boost the household *screening* non-response weights were applied to the dataset prior to the modelling of the household *cooperation* non-response,

whilst the product of the screening non-response weight and the address selection weight²⁹ were applied prior to the modelling of the non-response for the Muslim boost.³⁰

- 7.3.11 The variables used to model screening non-response at the address were also used to model household refusal. As before, the variables were selected to be consistent with previous years of the study. These were: Government Office Region, ACORN group (16 categories) and quintiles of the proportion of the output area's population belonging to a non-white ethnic minority group, based on data from the 2001 Census. A refusal weight (w3) was generated as the inverse of the saved predicted probabilities.
- 7.3.12 The model for Quarter 4 indicates responses tended to be higher in Wales and Yorkshire and Humberside relative to other regions, in ACORN groups P (High Rise Hardship) and N (Struggling Families) and in areas with a higher density ethnic minority population. The full model for Quarter 4 is presented in Table 7.11.

Table 7.11: Quarter 4 ethnic minority group sample household (cooperation) non-response model

	B	S.E.	Wald	df	Sig.	Exp(B)
Government Office Region			16.53	7	0.021	
East Midlands					(baseline)	
East of England	-0.43	0.31	1.91	1	0.167	0.65
London	-0.13	0.20	0.46	1	0.498	0.87
North West	-0.01	0.24	0.00	1	0.951	0.99
South East	-0.45	0.28	2.60	1	0.107	0.64
Wales	0.75	0.52	2.11	1	0.147	2.11
West Midlands	-0.08	0.22	0.12	1	0.735	0.93
Yorkshire and The Humber	0.45	0.27	2.83	1	0.093	1.57
Acorn Group			28.49	15	0.019	
A/B. Wealthy Executives/ Affluent Greys					(baseline)	
C. Flourishing Families	0.12	0.53	0.05	1	0.822	1.13
D. Prosperous Professionals	0.37	0.57	0.43	1	0.512	1.45
E. Educated Urbanites	0.32	0.37	0.73	1	0.393	1.37

²⁹ An address selection weight was used prior to modelling for the Muslim boost sample due to the unequal address selection probabilities between the two stratum; 2.5% to <10% and 10%+ Muslim population in output area. .

³⁰ For more information on why we weighted the data before modeling for non-response see Robert M. Groves and Mick P. Couper, *Non-response in Household Interview Surveys*, chapter 11 section 9.

Table 7.11: Quarter 4 ethnic minority group sample household (cooperation) non-response model (continued)

	B	S.E.	Wald	df	Sig.	Exp(B)
F. Aspiring Singles	0.84	0.37	5.15	1	0.023	2.32
G. Starting Out	-0.02	0.47	0.00	1	0.972	0.98
H. Secure Families	0.63	0.37	2.87	1	0.090	1.87
I/J. Settled Suburbia/Prudent Pensioners	0.84	0.51	2.71	1	0.100	2.31
K. Asian Communities	0.57	0.37	2.36	1	0.125	1.77
L. Post-Industrial Families	0.52	0.41	1.63	1	0.202	1.69
M. Blue-collar Roots	0.67	0.41	2.60	1	0.107	1.95
N. Struggling Families	0.95	0.38	6.34	1	0.012	2.59
O. Burdened Singles	0.43	0.43	1.02	1	0.312	1.54
P. High-Rise Hardship	1.22	0.51	5.76	1	0.016	3.40
Q. Inner City Adversity	0.66	0.37	3.14	1	0.077	1.93
Unclassified/missing	-0.06	1.21	0.00	1	0.961	0.94
% non-white ethnic minority population by Output Area (Census 2001)			13.03	4	0.011	
Bottom Quintile					(baseline)	
Second Quintile	-0.06	0.21	0.08	1	0.782	0.94
Third Quintile	0.37	0.20	3.51	1	0.061	1.45
Fourth Quintile	0.33	0.20	2.79	1	0.095	1.40
Top Quintile	0.50	0.21	5.55	1	0.019	1.65
Constant	-0.71	0.43	2.77	1	0.096	0.49

Notes:

1. The response is 1 = household response, 0 = no household response.
2. The model $R^2 = 0.034$ (Cox and Snells).
3. B is the estimate coefficient with standard error S.E.
4. The Wald-test measures the impact of the categorical variable on the model with the appropriate number of degrees of freedom df. If the test is significant ($\text{sig} < 0.05$) then the categorical variable is considered to be 'significantly associated' with the response variable.
5. The Wald test for each level of the categorical variable is also shown. This tests the difference between that level and the baseline category.

Dwelling unit selection weight

7.3.13 At each contacted address, the interviewer established the number of dwelling units. There are multiple dwelling units at a small proportion of addresses (<2%). In such cases the interviewer used a Kish grid to select a single dwelling unit at random to be included in the survey. The dwelling unit selection weight (w4) is

equivalent to the number of dwelling units at the selected address, and has been trimmed to a maximum of four.

Final calibrated household weight

- 7.3.14 The household weight for both boost samples is the product of the household screening non-response weight, the address selection weight, the household (cooperation) non-response weight and the dwelling unit selection weight ($w_1 \times w_2 \times w_3 \times w_4$). This weight is then calibrated to the population of England and Wales according to the 2008 mid-year household population estimates, using the same approach as applied to the core sample household weight.
- 7.3.15 The population estimates used for the calibration were age/sex (16 categories) and Government Office Unit (ten categories, including Wales) – see Tables 7.3 and 7.4. The final household weight for the combined sample across all 4 quarters (WtCombHH) is summarised in Table 7.12 This weight (WtCombHH) should be applied in the software package when interrogating the combined dataset for any household level estimates e.g. proportion of Black African single parent households in England and Wales.

Table 7.12: Quarter 1 to quarter 4 summary of final household weight (combined sample)

WtCHH	Number	Range	Minimum and maximum	Mean	Median	5th and 95th percentile
Household characteristics						
North East	619	14.3	0.04 – 14.37	1.52	1.46	0.07 – 2.85
North West	1,939	12.9	0.02 – 12.9	1.02	1.24	0.04 – 2.02
Yorkshire and The Humber	1,532	18.8	0.03 – 18.8	1.13	1.36	0.05 – 2.51
East Midlands	1,104	7.5	0.03 – 7.54	1.10	1.23	0.07 – 2.25
West Midlands	1,871	10.3	0.03 – 10.28	0.92	0.31	0.06 – 2.42
East	1,180	12.4	0.05 – 12.46	1.59	1.67	0.12 – 2.98
London	4,495	12.5	0.02 – 12.54	0.46	0.15	0.05 – 2.41
South East	1,717	10.1	0.00 – 10.15	1.16	1.37	0.08 – 2.19
South West	1,029	7.3	0.04 – 7.35	1.46	1.48	0.23 – 2.21
Wales	654	6.7	0.04 – 6.69	1.73	1.60	0.08 – 3.15
All	16,140	18.80	0.00 – 18.8	1.00	0.97	0.05 – 2.51

Calculation of individual weight for combined sample

7.3.16 This is made up of the household screening non-response weight, the address selection weight, the household (cooperation) non-response weight, the dwelling unit selection weight and the individual selection weight.

Individual selection weight

7.3.17 At each selected dwelling unit one individual was selected at random from all the eligible adults in the household. For core addresses this was any individual in the household aged 16 or over. For the ethnic minority boost this was any individual in the household aged 16 or over who was from an eligible ethnic group. For the Muslim boost this was any individual in the household aged 16 or over who was Muslim. The individual selection weights (w_5) are equivalent to the number of eligible individuals in the household. Without these weights, individuals in larger households would be under-represented in the sample. This weight was trimmed to a maximum of 5 for the boost samples.

Final calibrated individual weight

7.3.18 The individual weight for both boost samples is the product of the household screening non-response weight, the address selection weight, the household (cooperation) non-response weight, the dwelling unit selection weight and the individual selection weight ($w_1 \times w_2 \times w_3 \times w_4 \times w_5$).

7.3.19 This weight was then calibrated to the population of England and Wales aged 16 or over according to the 2008 mid-year household population estimates. The approach is that used for the core sample individual weight. The population estimates used for the calibration were age/sex (14 categories) and Government Office Region (10 categories, including Wales). The population figures are presented in Tables 7.6 and 7.7. The final individual weight for the combined sample ($WtFlnds$) is summarised in Table 7.13. This weight ($WtFlnds$) should be applied in the software package when interrogating the combined dataset for any individual level estimates e.g. proportion of Black African adults 16+ in England and Wales with one or more children.

Table 7.13: Quarter 4 summary of final individual weight (combined sample)

WtFinds	Number	Range	Minimum and maximum	Mean	Median	5th and 95th percentile
Individual Characteristics						
Region						
North East	619	8.6	0.02 – 8.62	1.25	1.23	0.07 – 2.68
North West	1,939	6.2	0.01 – 6.24	1.05	0.85	0.04 – 2.85
Yorkshire and The Humber	1,532	10.4	0.01 – 10.41	1.01	0.81	0.03 – 2.84
East Midlands	1,104	10.7	0.03 – 10.69	1.20	1.04	0.07 – 2.91
West Midlands	1,871	6.2	0.02 – 6.19	0.85	0.31	0.04 – 2.72
East	1,180	6.0	0.02 – 6.04	1.43	1.42	0.11 – 3.46
London	4,495	10.1	0.02 – 10.12	0.50	0.15	0.04 – 2.65
South East	1,717	9.4	0.02 – 9.42	1.44	1.47	0.12 – 3.38
South West	1,029	11.3	0.03 – 11.36	1.52	1.50	0.17 – 3.4
Wales	654	5.0	0.02 – 5.05	1.36	1.29	0.07 – 2.99
Sex						
Male	7,567	11.3	0.01 – 11.36	1.04	0.67	0.05 – 3.26
Female	8,573	10.1	0.01 – 10.12	0.97	0.78	0.04 – 2.73
Age group						
16-24	1,879	11.3	0.02 – 11.36	1.28	0.35	0.06 – 4.64
25-34	3,249	9.8	0.02 – 9.83	0.79	0.21	0.04 – 2.81
35-44	3,457	9.4	0.01 – 9.42	0.86	0.31	0.04 – 2.58
45-54	2,489	7.6	0.01 – 7.62	1.05	0.83	0.05 – 2.92
55-64	2,053	10.7	0.01 – 10.69	1.15	1.26	0.06 – 2.52
65-74	1,661	4.5	0.02 – 4.49	1.00	0.88	0.06 – 1.97
75+	1,352	4.1	0.02 – 4.16	1.15	1.03	0.08 – 2.26
All	16,140	11.3	0.01 – 11.36	1.00	0.75	0.05 – 2.92

Quarterly adjustment

7.3.20 As in the core sample the quarterly weights were used in the generation of weights for the yearly data. To avoid over-representing any quarter in the final yearly combined household and individual weights, a quarterly adjustment was made to make the yearly sample even across quarters.

Section 8

Standard errors

8.1 Introduction

- 8.1.1 Tables in this chapter present estimates for standard errors for key variables discussed in the topic reports, taking into account the complex sample design of the survey.

8.2 Sources of error in surveys

- 8.2.1 Survey results are subject to various sources of error. The total error in a survey estimate is the difference between the estimate derived from the data collected and the true value for the population. The total error can be divided into two main types: systematic and random error.

Systematic error

- 8.2.2 Systematic error, or bias, covers those sources of error which will not average to zero over repeats of the survey. Bias may occur, for example, if a certain section of the population is excluded from the sampling frame, because non-respondents to the survey have different characteristics to respondents, or if interviewers systematically influence responses in one way or another. Substantial efforts have been made to avoid systematic errors.

Random error

- 8.2.3 An important component of random error is sampling error, which is the error that arises because the estimate is based on a random sample rather than a full census of the population. The results obtained for any single sample may, by chance, vary from the true values for the population but the variation would be expected to average to zero over a number of repeats of the survey. The amount of variation depends on both the size of the sample and the sample design.
- 8.2.4 Random error may also result from other sources such as variations in respondents' interpretation of the questions, or variations in the way different interviewers ask questions. Efforts are made to minimise these effects through pilot work and interviewer training. The impact of this random variation is reflected in the standard errors presented here.

8.3 Standard errors for complex sample designs

- 8.3.1 The Citizenship Survey uses a two-stage stratified sample design. In considering the reliability of estimates, standard errors calculated on the basis of a simple random sample design will not reflect the true variation because of the complex sample design. The two-stage sample of addresses can lead to a substantial increase in standard error if the households or individuals within primary sampling units are relatively homogenous but the primary sampling units differ from one another.
- 8.3.2 Stratification tends to reduce standard error and is of most advantage where the stratification factor is related to the characteristics of interest on the survey.
- 8.3.3 In a complex sample design, the size of the standard error depends on how the characteristic of interest is spread within and between the primary sampling units and strata, and this is taken into account in the way data are grouped in order to calculate the standard error. For the Citizenship Survey, the weighting for different sampling probabilities (i.e. the boost samples and the sub-sampling of adults within households) and different response rates also increases the size of the standard errors compared with an equal probability sample of the same size.
- 8.3.4 The method for calculating standard error compares the differences between totals for adjacent primary sampling units (wards/output areas) in the characteristic of interest. The ordering of primary sampling units reflects the ranking of wards/output areas on the stratifiers used in the sample design.

8.4 Design factor (deft)

- 8.4.1 The design factor, or deft, is the ratio of the standard error of an estimate to the standard error that would have resulted had the survey design been a simple random sample of the same size. The size of the design factor varies between survey variables according to the degree to which a characteristic is clustered within primary sampling units, or is distributed between strata, and the impact of the weighting. For a single variable, the size of the factor also varies according to the size of the subgroup on which the estimate is based, and on the distribution of the subgroup between primary sampling units and strata. Design factors below 1.0 show that the complex sample design improved on the estimate that would have expected from a simple random sample, probably due to the benefits of stratification. Design factors greater than 1.0 show less reliable estimates than might be gained from a simple random sample, due to the effects of clustering and weighting.

- 8.4.2 The standard error and defts for selected survey estimates are shown in tables 8.1 to 8.14. These can be used to estimate likely sampling errors for other variables on the basis of their similarity to one of the variables presented. The standard error (se) of a proportion (p) based on a simple random sample (srs) multiplied by the deft gives the standard error of a complex design.

$$se(p) = deft \times se(p)_{srs}$$

$$se(p)_{srs} = \sqrt{\frac{p(100-p)}{n}} \quad 31$$

- 8.4.3 The formula to calculate the standard error of the difference between two percentages for a complex sample design is:

$$se(p_1 - p_2) = \sqrt{\frac{deft_1^2(p_1(100-p_1))}{n_1} + \frac{deft_2^2(p_2(100-p_2))}{n_2}}$$

- 8.4.4 Where p_1 and p_2 are observed percentages for the two subsamples and n_1 and n_2 are the subsample sizes.

8.5 Confidence intervals

- 8.5.1 The estimate produced from a sample survey will rarely be identical to the population value, but statistical theory allows us to measure the accuracy of any survey result. The standard error can be estimated from the values obtained for the sample and allows the calculation of confidence intervals, which indicate the range of random variation in the survey estimates.
- 8.5.2 It is common, when quoting confidence intervals, to refer to the 95 per cent confidence interval around a survey estimate. This is calculated at 1.96 times the standard error on either side of the estimated percentage or mean since, under a normal distribution, 95 per cent of values lie within 1.96 standard errors of the mean value. If it were possible to repeat the survey under the same conditions many times, 95 per cent of these confidence intervals would contain the population values.
- 8.5.3 The 95 per cent confidence interval for the difference between two percentages is then given by:

$$(p_1 - p_2) \pm 1.96 \times se(p_1 - p_2)$$

³¹ The precise formula uses $n-1$ as the denominator but this equates to n in large samples.

- 8.5.4 If this confidence interval includes zero then the hypothesis that the two proportions are the same and the observed difference is due to chance alone is not rejected. If the interval does not include zero then it is unlikely (less than five per cent probability) that the observed difference could have occurred by chance and this constitutes a 'significant difference' at the 95 per cent confidence level.
- 8.5.5 The 95 per cent confidence level was used for all significance testing in the analysis which is reported in the topic reports on the survey.

8.6 Standard errors for the 2009–10 Citizenship Survey

- 8.6.1 The standard errors were calculated on weighted data using SAS.³² As mentioned earlier, weighting for different sampling probabilities and different response rates results in larger sampling errors than for an equal-probability sample without weights. However, using population totals to control for differential non-response tends to lead to a small reduction in the errors. The method used to calculate the sampling errors correctly allows for the inflation in the sampling errors caused by the first type of weighting but, in treating the second type of weighting in the same way as the first, incorrectly inflates the estimates further. Therefore the standard errors and defts presented are likely to be slight over-estimates. Weighted data were used so that the values of the percentages and means were the same as those in the substantive chapters of the topic reports for 2009–10.
- 8.6.2 Tables 8.1 to 8.14 show the standard error and defts for selected survey estimates.
- 8.6.3 For most measures, the sampling errors were based on the core sample as this provides more robust estimates than the combined core and boost samples.
- 8.6.4 Sampling errors for estimates for ethnic minority subgroups were based on the combined sample.

³² SAS is a statistical analysis software package. For further details of the method of calculation see: <http://www2.sas.com/proceedings/sugi27/p263-27.pdf>

Table 8.1: Sampling errors for weighted core sample data: Political efficacy, England only

Characteristic	Population	Percentage (p)	Unweighted base	Standard error of p	Design factor (Deft)
		percentage	number	percentage	number
Political efficacy					
Percentages who agreed that:					
They could influence decisions affecting their local area	Respondents living in England	37.0	8,307	0.0064	1.20
They could influence decisions affecting Britain	Respondents living in England	20.1	8,402	0.0057	1.31

Table 8.2: Sampling errors for weighted core sample data: Labour market discrimination, all respondents

Characteristic	Population	Percentage (p)	Unweighted base	Standard error of p	Design factor (Deft)
		percentage	number	percentage	Number
Racial prejudice and discrimination					
Percentages who:					
Had been discriminated against when refused a job in the last five years	Respondents who were working as employees or who had looked for a job in the past five years	6.84	5,776	0.0039	1.17
Had been discriminated against when refused a promotion/move to a better position in the last five years	Respondents who were currently working as employees	5.78	5,619	0.0035	1.14

Table 8.3: Sampling errors for weighted core sample data: Views on the neighbourhood, England

Characteristic	Population	Percentage (p)	Unweighted base	Standard error of p	Design factor (Deft)
		<i>percentage</i>	<i>number</i>	<i>percentage</i>	<i>number</i>
Views on the neighbourhood					
Percentages who:					
Felt they belonged very strongly to the neighbourhood	All respondents	35.63	8,653	0.0065	1.25

Table 8.4: Sampling errors for weighted core sample data: Active participation in communities, England

Characteristic	Population	Percentage (p)	Unweighted base	Standard error of p	Design factor (Deft)
		<i>percentage</i>	<i>number</i>	<i>percentage</i>	<i>number</i>
Active participation in communities					
Percentages who:					
Had undertaken any civic engagement activity or formal volunteering at least once in the previous 12 months	Respondents in England	59.29	8,712	0.0068	1.30
Had undertaken any civic consultation at least once a month in the previous 12 months	Respondents in England	18.46	8,712	.0048	1.12

Table 8.4: Sampling errors for weighted core sample data: Active participation in communities, England (continued)

Characteristic	Population	Percentage (p)	Unweighted base	Standard error of p	Design factor (Deft)
		<i>percentage</i>	<i>number</i>	<i>percentage</i>	<i>number</i>
Had undertaken civic participation at least once in the previous 12 months	Respondents in England	34.19	8,712	.0061	1.21
Had undertaken civic activism activities at least once in the previous 12 months	Respondents in England	10.20	8,712	.0038	1.18
Had undertaken civic participation at least once a month in the previous 12 months	Respondents in England	2.52	8,712	.0019	1.12
Participated in informal volunteering at least once in the previous 12 months	Respondents in England	54.20	8,712	.0072	1.34
Participated in informal volunteering at least once a month in the previous 12 months	Respondents in England	29.31	8,712	.0059	1.20
Participated in formal volunteering at least once in the previous 12 months	Respondents in England	40.29	8,712	.0069	1.31
Participated in formal volunteering at least once a month in the previous 12 months	Respondents in England	25.07	8,712	.0057	1.22

Table 8.5: Sampling errors for weighted core sample data: Racial discrimination by organisations, all respondents					
Characteristic	Percentage (p)	Unweighted base	Standard error of p	Design factor (Def't)	
	<i>percentage</i>	<i>number</i>	<i>percentage</i>	<i>number</i>	
Racial prejudice					
<i>Percentages who expected organisations to treat them worse than people of other races:</i>					
A council housing dept. or housing association					
treated worse than others	17.70	9,304	0.005		1.26
A local school					
treated worse than others	2.69	9,304	0.0019		1.15
A local doctors surgery					
treated worse than others	1.76	9,304	0.0016		1.16
The police					
treated worse than others	6.03	9,304	0.0033		1.33
The Prison Service					
treated worse than others	3.35	9,304	0.0027		1.42
The courts					
treated worse than others	5.14	9,304	0.0031		1.35
The Crown Prosecution Service					
treated worse than others	4.80	9,304	0.003		1.33
The Probation Service					
treated worse than others	3.11	9,304	0.003		1.44

Table 8.6: Sampling errors for weighted core sample data: Community cohesion in the local area, England only

Characteristic	Population	Percentage (p)	Unweighted base	Standard error of p	Design factor (Deft)
		<i>percentage</i>	<i>number</i>	<i>percentage</i>	<i>Number</i>
Community cohesion in the local area					
Percentages who agreed that:					
The local area is a place where people from different backgrounds get on well together	<i>All respondents</i>	84.61	7,403	.0052	1.23

Table 8.7: Sampling errors for weighted core sample data: Meaningful interactions, England

Characteristic	Population	Percentage (p)	Unweighted base	Standard error of p	Design factor (Deft)
		<i>percentage</i>	<i>number</i>	<i>percentage</i>	<i>Number</i>
Meaningful interaction					
Percentages who mix socially with people from different ethnic or religious groups:					
Mix socially	<i>All respondents</i>	79.73	8,704	.006	1.39

Table 8.8: Sampling errors for weighted combined sample data: Political efficacy by ethnicity, England only

Characteristic	Population	Percentage (p)	Unweighted base	Standard error of p	Design factor (Deft)
		<i>percentage</i>	<i>number</i>	<i>percentage</i>	<i>number</i>
Political efficacy					
Percentages who tend to agree that:					
They could influence decisions affecting their local area	<i>Respondents in England</i>				
	White	36.35	7,687	0.0067	1.21
	All Asian	44.63	3,710	0.0125	1.53
	Indian	42.33	1,186	0.0187	1.31
	Pakistani	46.71	1,539	0.0207	1.63
	Bangladeshi	43.82	577	0.0366	1.77
	All Black	49.22	1,863	0.0184	1.58
	Caribbean	43.20	856	0.0252	1.49
	African	54.15	965	0.0245	1.52
	Mixed race	34.76	420	0.0329	1.42
	Chinese/Other	40.99	674	0.0314	1.66
	All ethnic minority groups	44.70	6,667	0.0100	1.64

Table 8.8: Sampling errors for weighted combined sample data: Political efficacy by ethnicity, England only (continued)

Characteristic	Population	Percentage (p)	Unweighted base	Standard error of p	Design factor (Deft)
		percentage	number	percentage	number
They could influence decisions affecting Britain	<i>Respondents in England</i>				
	White	18.67	7,785	0.0056	1.27
	All Asian	33.01	3,678	0.0129	1.66
	Indian	31.35	1,189	0.0188	1.40
	Pakistani	34.66	1,505	0.0212	1.73
	Bangladeshi	30.82	573	0.0327	1.70
	All Black	34.42	1,885	0.0180	1.65
	Caribbean	26.08	862	0.0331	2.21
	African	41.49	977	0.0230	1.46
	Mixed race	26.20	419	0.0331	1.54
	Chinese/Other	25.86	678	0.0258	1.54
	All ethnic minority groups	31.99	6,660	0.0100	1.74

Table 8.9: Sampling errors for weighted combined sample data: Labour market discrimination by ethnicity, all respondents

Characteristic	Population	Percentage (p)	Unweighted base	Standard error of p	Design factor (Deft)
		<i>percentage</i>	<i>number</i>	<i>percentage</i>	<i>number</i>
Racial prejudice and discrimination					
Percentages who:					
Had been refused a job in the last five years	<i>Respondents who were working as employees or who had looked for a job in the past 5 years</i>				
	White	5.88	5,280	0.0037	1.15
	All Asian	9.66	2,566	0.0076	1.31
	Indian	7.96	879	0.0098	1.08
	Pakistani	12.26	977	0.0172	1.63
	Bangladeshi	10.34	392	0.0200	1.30
	All Black	16.59	1,441	0.0140	1.43
	Caribbean	15.95	640	0.0231	1.60
	African	17.14	763	0.0170	1.24
	Mixed race	14.93	341	0.0289	1.49
	Chinese/Other	7.56	501	0.0138	1.17
	All ethnic minority groups	11.83	4,849	0.0063	1.36

Table 8.9: Sampling errors for weighted combined sample data: Labour market discrimination by ethnicity, all respondents (continued)

Characteristic	Population	Percentage (p)	Unweighted base	Standard error of p	Design factor (Deft)
		<i>percentage</i>	<i>number</i>	<i>percentage</i>	<i>number</i>
Had been refused a promotion/ move to a better position in the last five years	<i>Respondents who were currently working as employees</i>				
	White	5.29	5,144	0.0036	1.14
	All Asian	7.67	2,422	0.0068	1.26
	Indian	9.16	841	0.0115	1.16
	Pakistani	6.26	917	0.0095	1.19
	Bangladeshi	7.16	359	0.0192	1.41
	All Black	14.68	1,341	0.0141	1.45
	Caribbean	15.39	595	0.0211	1.43
	African	14.53	709	0.0202	1.52
	Mixed race	7.19	315	0.0158	1.08
	Chinese/Other	6.24	475	0.0120	1.08
	All ethnic minority groups	9.45	4,553	0.0057	1.31

Table 8.10: Sampling errors for weighted combined sample data: Views on the neighbourhood by ethnicity, England						
Characteristic	Population	Percentage (p)	Unweighted base	Standard error of p	Design factor (Deft)	
		<i>percentage</i>	<i>number</i>	<i>percentage</i>	<i>number</i>	
Felt they belonged strongly to the neighbourhood	<i>All respondents</i>					
	White	76.60	7,982	0.0057	1.19	
	All Asian	81.11	4,135	0.0101	1.65	
	Indian	80.64	1,310	0.0152	1.39	
	Pakistani	82.91	1,713	0.0167	1.83	
	Bangladeshi	87.74	656	0.0187	1.46	
	All Black	74.64	2,014	0.0135	1.39	
	Caribbean	76.04	917	0.0220	1.56	
	African	73.58	1,047	0.0182	1.34	
	Mixed race	69.48	442	0.0364	1.66	
	Chinese/Other	71.58	780	0.0248	1.54	
	All ethnic minority groups	77.24	7,371	0.0082	1.67	

Table 8.11: Sampling errors for weighted combined sample data: Active participation in communities by ethnicity, England

Characteristic	Population	Percentage (p)	Unweighted base	Standard error of p	Design factor (Deft)
		percentage	number	percentage	number
Active participation in communities					
Percentages who:					
Had undertaken any civic engagement or formal volunteering activity at least once in the previous 12 months	All respondents				
	White	60.57	8,032	0.0072	1.32
	All Asian	48.90	4,168	0.0129	1.66
	Indian	49.37	1,317	0.0191	1.39
	Pakistani	47.06	1,727	0.0214	1.78
	Bangladeshi	53.33	664	0.0335	1.73
	All Black	55.03	2,045	0.0170	1.54
	Caribbean	55.98	924	0.0263	1.61
	African	54.10	1,071	0.0239	1.57
	Mixed race	57.42	449	0.0355	1.52
	Chinese/Other	38.49	788	0.0294	1.70
	All ethnic minority groups	49.79	7,450	0.0101	1.74

Table 8.11: Sampling errors for weighted combined sample data: Active participation in communities by ethnicity, England (continued)

Characteristic	Population	Percentage (p)	Unweighted base	Standard error of p	Design factor (Deft)
		percentage	number	percentage	number
Had undertaken civic participation at least once in the previous 12 months	All respondents				
	White	35.56	8,032	0.0065	1.21
	All Asian	25.74	4,168	0.0111	1.64
	Indian	23.46	1,317	0.0162	1.38
	Pakistani	24.11	1,727	0.0154	1.50
	Bangladeshi	34.28	664	0.0391	2.12
	All Black	24.07	2,045	0.0137	1.44
	Caribbean	26.73	924	0.0208	1.43
	African	21.96	1,071	0.0180	1.42
	Mixed race	31.30	449	0.0359	1.64
	Chinese/Other	18.45	788	0.0192	1.39
	All ethnic minority groups	24.72	7,450	0.0082	1.64

Table 8.11: Sampling errors for weighted combined sample data: Active participation in communities by ethnicity, England
(continued)

Characteristic	Population	Percentage (p)	Unweighted base	Standard error of p	Design factor (Deft)
		<i>percentage</i>	<i>number</i>	<i>percentage</i>	<i>number</i>
Participated in informal volunteering at least once in the previous 12 months	<i>All respondents</i>				
	White	55.38	8,032	0.0075	1.35
	All Asian	44.62	4,168	0.0127	1.65
	Indian	46.71	1,317	0.0182	1.32
	Pakistani	42.91	1,727	0.0219	1.84
	Bangladeshi	39.31	664	0.0349	1.84
	All Black	51.79	2,045	0.0176	1.59
	Caribbean	49.83	924	0.0285	1.73
	African	53.04	1,071	0.0241	1.58
	Mixed race	56.85	449	0.0348	1.49
	Chinese/Other	38.44	788	0.0324	1.87
	All ethnic minority groups	46.63	7,450	0.0104	1.79

Table 8.11: Sampling errors for weighted combined sample data: Active participation in communities by ethnicity, England (continued)

Characteristic	Population	Percentage (p)	Unweighted base	Standard error of p	Design factor (Deft)
		<i>percentage</i>	<i>number</i>	<i>percentage</i>	<i>number</i>
Participated in informal volunteering at least once a month in previous 12 months	<i>All respondents</i>				
	White	29.71	8,032	0.0063	1.23
	All Asian	22.87	4,168	0.0108	1.66
	Indian	23.36	1,317	0.0178	1.53
	Pakistani	23.28	1,727	0.0159	1.56
	Bangladeshi	20.28	664	0.0256	1.64
	All Black	30.60	2,045	0.0170	1.67
	Caribbean	29.55	924	0.0258	1.72
	African	30.89	1,071	0.0230	1.63
	Mixed race	31.85	449	0.0341	1.55
	Chinese/Other	17.43	788	0.0193	1.43
	All ethnic minority groups	24.89	7,450	0.0087	1.73

Table 8.11: Sampling errors for weighted combined sample data: Active participation in communities by ethnicity, England
(continued)

Characteristic	Population	Percentage (p)	Unweighted base	Standard error of p	Design factor (Deft)
		<i>percentage</i>	<i>number</i>	<i>percentage</i>	<i>number</i>
Participated in formal volunteering at least once in the previous 12 months	<i>All respondents</i>				
	White	41.20	8,032	0.0073	1.33
	All Asian	29.43	4,168	0.0117	1.66
	Indian	32.02	1,317	0.0182	1.41
	Pakistani	26.42	1,727	0.0186	1.76
	Bangladeshi	27.92	664	0.0303	1.74
	All Black	38.64	2,045	0.0181	1.68
	Caribbean	37.79	924	0.0291	1.82
	African	38.77	1,071	0.0236	1.58
	Mixed race	41.45	449	0.0374	1.61
	Chinese/Other	24.92	788	0.0233	1.51
	All ethnic minority groups	32.20	7,450	0.0098	1.80

Table 8.11: Sampling errors for weighted combined sample data: Active participation in communities by ethnicity, England (continued)

Characteristic	Population	Percentage (p)	Unweighted base	Standard error of p	Design factor (Deft)
		<i>percentage</i>	<i>number</i>	<i>percentage</i>	<i>number</i>
Participated in formal volunteering at least once a month in the previous 12 months	All respondents				
	White	25.75	8,032	0.0060	1.23
	All Asian	15.89	4,168	0.0086	1.53
	Indian	16.33	1,317	0.0131	1.29
	Pakistani	15.15	1,727	0.0145	1.68
	Bangladeshi	16.29	664	0.0247	1.73
	All Black	24.68	2,045	0.0162	1.70
	Caribbean	24.83	924	0.0269	1.89
	African	23.88	1,071	0.0194	1.49
	Mixed race	22.98	449	0.0275	1.38
	Chinese/Other	13.20	788	0.0176	1.46
	All ethnic minority groups	18.43	7,450	0.0074	1.65

Table 8.12: Sampling errors for weighted combined sample data: Racial discrimination by organisations by ethnicity, all respondents

Characteristic	Population	Percentage (p)	Unweighted base	Standard error of p	Design factor (Deft)
		<i>percentage</i>	<i>number</i>	<i>percentage</i>	<i>number</i>
Racial prejudice					
<i>Percentages who expected organisations to treat them worse than people of other races:</i>					
A council housing dept. or housing association	<i>All respondents</i>				
treated worse than others	White	19.01	8,611	0.0055	1.31
	All ethnic minority groups	6.98	7,516	0.0049	1.68
A local school	<i>All respondents</i>				
treated worse than others	White	2.59	8,611	0.0020	1.17
	All ethnic minority groups	2.69	7,516	0.0023	1.24
A local doctors surgery	<i>All respondents</i>				
treated worse than others	White	1.73	8,611	0.0017	1.22
	All ethnic minority groups	1.67	7,516	0.0019	1.26

Table 8.12: Sampling errors for weighted combined sample data: Racial discrimination by organisations by ethnicity, all respondents (continued)

Characteristic	Population	Percentage (p)	Unweighted base	Standard error of p	Design factor (Deft)
		<i>percentage</i>	<i>number</i>	<i>percentage</i>	<i>number</i>
The Police	<i>All respondents</i>				
treated worse than others	White	4.60	8,611	0.0027	1.20
	All ethnic minority groups	14.67	7,516	0.0078	1.91
The Prison Service	<i>All respondents</i>				
treated worse than others	White	2.38	8,611	0.0021	1.29
	All ethnic minority groups	9.09	7,516	0.0067	2.01
The courts	<i>All respondents</i>				
treated worse than others	White	4.55	8,611	0.0028	1.25
	All ethnic minority groups	7.39	7,516	0.0057	1.90
The Crown Prosecution Service	<i>All respondents</i>				
treated worse than others	White	4.21	8,611	0.0027	1.23
	All ethnic minority groups	7.14	7,516	0.0056	1.90
The Probation Service	<i>All respondents</i>				
treated worse than others	White	2.44	8,611	0.0021	1.29
	All ethnic minority groups	6.17	7,516	0.0053	1.89

Table 8.13: Sampling errors for weighted combined sample data: Community cohesion in the local area by ethnicity, England only

Characteristic	Population	Percentage (p)	Unweighted base	Standard error of p	Design factor (Deft)
		<i>percentage</i>	<i>number</i>	<i>percentage</i>	<i>number</i>
Community cohesion in the local area					
Percentages who agreed that:					
The local area is a place where people from different backgrounds get on well together	<i>All respondents</i>				
	White	84.28	6,761	0.0056	1.26
	All Asian	88.98	3,991	0.0075	1.51
	Indian	88.27	1,277	0.0136	1.51
	Pakistani	89.81	1,648	0.0115	1.54
	Bangladeshi	90.19	632	0.0153	1.30
	All Black	86.93	1,918	0.0117	1.52
	Caribbean	88.38	873	0.0154	1.41
	African	86.34	998	0.0173	1.59
	Mixed race	80.43	426	0.0282	1.47
	Chinese/Other	88.90	732	0.0151	1.30
	All ethnic minority groups	87.79	7,067	0.0057	1.47

Table 8.14: Sampling errors for weighted combined sample data: Meaningful interactions by ethnicity, England

Characteristic	Population	Percentage (p)	Unweighted base	Standard error of p	Design factor (Deft)
		<i>percentage</i>	<i>number</i>	<i>percentage</i>	<i>number</i>
Meaningful interaction					
<i>Percentages who mix socially with people from other ethnic or religious groups:</i>					
Had meaningful interactions	<i>All respondents</i>				
	White	78.02	8,022	0.0065	1.40
	All Asian	94.36	4,161	0.0071	2.00
	Indian	95.06	1,315	0.0105	1.76
	Pakistani	93.20	1,723	0.0110	1.81
	Bangladeshi	91.73	663	0.0314	2.94
	All Black	97.17	2,038	0.0056	1.51
	Caribbean	97.41	919	0.0058	1.11
	African	96.89	1,069	0.0086	1.62
	Mixed race	99.23	449	0.0036	0.87
	Chinese/Other	95.90	785	0.0091	1.29
	All ethnic minority groups	95.68	7,433	0.0046	1.93

Section 9

Data User Guide

9.1 Introduction

- 9.1.1 This chapter provides a user guide for those conducting analysis of the Citizenship Survey dataset. The dataset is available on the UK Data Archive³³ in SPSS format, and the guide assumes that analysis will be conducted in SPSS.

9.2 Selecting cases for analysis

Core and boost samples

- 9.2.1 The dataset contains data from the core, the ethnic minority boost samples and the Muslim boost sample. The sample can be selected using the variable 'samptype' where 1=core, 2 or 3 = ethnic minority boost sample and 4 = faith boost sample.
- 9.2.2 For example, under 'data', select 'samptype=1' for the core sample, and then run crosstabs, frequencies, etc.
- 9.2.3 Most analysis should be conducted using the core sample only, which has a total unweighted base of 9,305. Analysis by ethnicity, religion or subgroups based on ethnicity such as religious group, religious activity or country of birth should use the combined core and boost samples which has an unweighted base of 16,140.

Quarters

- 9.2.4 The dataset contains data from fieldwork between 1 April 2009 and 31 March 2010, broken down into quarters. To perform analysis on an individual quarter use the variable 'Quarter' and select the appropriate quarter:
- Quarter 1: April to June 2009;
 - Quarter 2: July to September 2009;
 - Quarter 3: October to December 2009; and
 - Quarter 4: January to March 2010.
- 9.2.5 For example, under 'data', select quarter=1, and then run crosstabs, frequencies, etc.

³³ <http://www.data-archive.ac.uk/>

9.3 Variables

- 9.3.1 The dataset is ordered in the following way: first, variables containing serial number and type of sample; second, key demographic variables; third, survey question responses; fourth, further classificatory data on the respondent and household reference person ; fifth, derived variables; and sixth, weights. Any queries on published variables and requests for matched-in data need to be sent to DCLG in the first instance.
- 9.3.2 Variables are named similarly or exactly to match the question names in the questionnaire, with numbered suffixes when multiple variables in the dataset relate to the same question.
- 9.3.3 For further information about variable names please refer to the questionnaire (Annex E under separate cover to this report) and the list of key derived variables in Annex F (under separate cover to this report).

9.4 Multiple response questions

- 9.4.1 Where a respondent was able to give multiple answers to one question (a multiple response question), the question has been represented in the dataset by dummy variables, which are coded 'yes' or 'no' depending on whether the respondent chose this response or not. This aids analysis as it avoids the need to recode each multiple response question.

9.5 Missing values

- 9.5.1 For most variables, 'don't know' and 'refusal' responses have been set as missing values. However, for some variables 'don't know' is a valid response as per DCLG's specification.

9.6 Weighting

- 9.6.1 Five sets of weights are included in the dataset, allowing analysis by individual quarter or by the full year's data. Additionally, weights are included for analysis by individual or by household for both the core and full samples. Generally, analysis is carried out at the individual level on the core sample for the entire year, and in this instance the weight 'WtCInds' should be used.
- 9.6.2 The weight variable names and the description of the weights are listed below in Table 9.1.

Table 9.1: Weight variable names and descriptions

Weight	Description
WtFInds	Individual weight for combined sample for full survey year
WtCombHH	Household weight for combined sample for full survey year
WtCInds	Individual weight for core sample for full survey year
WtCHhds	Household weight for core sample for full survey year
Q1WtCIn	Quarter 1 Individual weight for core sample
Q1WtFln	Quarter 1 Individual weight for combined sample
Q1WtCBMEInd	Quarter 1 Individual weight for core and ethnic minority sample
Q1WtCHh	Quarter 1 Household weight for core sample
Q1WtFHh	Quarter 1 Household weight for combined sample
Q2WtCIn	Quarter 2 Individual weight for core sample
Q2WtFln	Quarter 2 Individual weight for combined sample
Q2WtCBMEInd	Quarter 2 Individual weight for core and ethnic minority sample
Q2WtCHh	Quarter 2 Household weight for core sample
Q2WtFHh	Quarter 2 Household weight for combined sample
Q3WtFln	Quarter 3 Individual weight for combined sample
Q3WtFHh	Quarter 3 Household weight for combined sample
Q3WtCIn	Quarter 3 Individual weight for core sample
Q3WtCHh	Quarter 3 Household weight for core sample
Q4WtFln	Quarter 4 Individual weight for combined sample
Q4WtFHh	Quarter 4 Household weight for combined sample
Q4WtCIn	Quarter 4 Individual weight for core sample
Q4WtCHh	Quarter 4 Household weight for core sample

9.7 Significance testing and standard errors

- 9.7.1 Significance tests for the topic reports 2009-10 have been carried out using the complex survey design features in SPSS and standard errors have been calculated in SPSS. The variables `primary sampling unit_scr` and `Strata_scr` indicate the primary sampling unit and strata of each case respectively.
- 9.7.2 Additionally, design factors were calculated for the quarterly data tables provided to DCLG, using SAS. Tables 8.1 to 8.14 show the standard error and defts for selected survey estimates for the 2009-10 survey.

Section 10

References

Cox, D.R and Snell, E.J (1989), *The Analysis of Binary Data*, Second Edition, (London: Chapman and Hall).

Deville, J.C. & Särndal, C.E. (1992). "Calibration estimators in survey sampling", *Journal of the American Statistical Association*, volume 87, pp. 376-382.

Green, Hazel, and Farmer, Christine at Office for National Statistics, (2004), *2003 Home Office Citizenship Survey: People, Families and Communities. Technical Report* (London: Home Office).

Groves, Robert M., and Couper, Mick P., (1998), *Nonresponse in Household Interview Surveys*, (Oxford: Wiley-Blackwell).

Kish, Leslie, (1949), 'A Procedure for Objective Respondent Selection Within the Household', *Journal of the American Statistical Association*, 44, pp. 380-87.

NatCen, (2008), *2008-9 Citizenship Survey: Technical Report*, (London: Department for Communities and Local Government).

Office for National Statistics, *Geography – Government Office Regions*. Accessible at: <http://www.statistics.gov.uk/geography/gor.asp>

Section 11

Annexes

The annexes to this technical report are under separate cover.

ISBN: 978-1-4098-2940-9

ISBN 978-1-4098-2940-9



9 781409 829409