

**British Crime Survey
2009-10 Technical report**

TNS-BMRB Report

Date October 2010

Content

1.	Background.....	1
1.1	Introduction to the British Crime Survey.....	1
1.2	Outputs from the BCS.....	3
1.3	Structure of the technical report.....	4
2.	Sample Design	5
2.1	Introduction.....	5
2.2	Target issued and achieved sample in Police Force Areas	5
2.3	A partially clustered sample	6
2.4	Sampling of addresses	6
2.5	Stratification.....	8
2.6	Allocation of sample to fieldwork quarters and months	8
2.7	Sampling of individuals	9
2.8	Extension to include 10-15 year olds	9
3.	Questionnaire Content and Development	11
3.1	Structure and coverage of the questionnaire	11
3.2	Life event calendar	28
3.3	Questionnaire development.....	28
3.4	Final questionnaire and revisions.....	29
3.5	Allocation of sample within CAPI	30
3.6	Features of Quancept used in the BCS	31
4.	Fieldwork	35
4.1	Briefing of interviewers	35
4.2	Supervision and quality control	36
4.3	Fieldwork dates and fieldwork management	37
4.4	Fieldwork procedures and documents	39
4.5	Presence of others during the interview	41
4.6	Length of interview	46
4.7	Response rate and reasons for non-response: core sample	51
4.8	Response to the self-completion questionnaire.....	59
4.9	Full and Partial Interviews.....	64
5.	Data Processing	65
5.1	Offence coding	65
5.2	Other coding.....	72
5.3	Coding of occupation and socio-economic classification.....	73

6.	Data Output	75
6.1	Introduction.....	75
6.2	Delivery of data output.....	75
6.3	Content of SPSS data file	76
6.4	Conventions used on SPSS Data Files.....	77
7.	Weighting.....	83
7.1	Overview of weighting.....	83
7.2	Component weights	84
7.3	Core sample weights	86
7.4	Calibration Weights	87
7.5	Analysis of non response.....	88
8.	Comparing key survey variables with the population.....	91
9.	The Under 16s survey.....	95
9.1	Background	95
9.2	10 to 15 year old sample	95
9.3	Questionnaire content and development.....	96
9.4	Fieldwork procedures and documents	101
9.5	Item non-response.....	102
9.6	Presence of others during the interview	102
9.7	Acceptance of self completion and use of Audio-CASI	104
9.8	Interview Length	105
9.9	Response rate and reasons for non response: 10-15 year old sample	106
9.10	Data processing.....	109
9.11	Home office coding.....	109
9.12	Data output.....	111
9.13	Weighting	111
	Annex 1 – Non-response analysis (based on 2008-09 survey).....	113
	Annex 2 – 2009-10 British Crime Survey: Computation of the under 16s non-response rate	139

1. Background

1.1 Introduction to the British Crime Survey

The British Crime Survey (BCS) is a well-established study and one of the largest social research surveys conducted in England and Wales. The survey was first conducted in 1982 and ran at roughly two yearly intervals until 2001, when it became a continuous survey¹. The survey is carried out for the Home Office, and is managed by a team of researchers in the Home Office Statistics Unit. They develop each survey in collaboration with an external research organisation. Since 2001 BMRB Social Research (now TNS-BMRB) has been the sole contractor for the survey.

Since the survey became continuous in 2001 there have been few significant changes to the design of the survey. Where changes have been incorporated these have been described in detail in the relevant technical reports. The most significant changes to the design of the survey have been:

- Increase of the core sample size to 46,000 to allow a target of at least 1,000 interviews in each Police Force Area (2004-05 technical report)
- Changes to the clustering of sample for interview (2008-09 technical report)
- Removal of the requirement for an additional boost of 4,000 interviews with non-white respondents
- Removal of the requirement for an additional boost of 2,000 interviews with respondents aged 16-24.
- Extension of the survey to cover young people aged 10-15 (2009-10 technical report)

In 2009-10 the total core sample size was the same as in the previous year, with approximately 46,000 core adult interviews being conducted across the year. The survey was designed to achieve a minimum of around 1,000 core interviews in each Police Force Area in England and Wales. The additional boost of approximately 2,000 interviews with young adults aged 16-24 was not included in 2009-10. From January 2009 the survey was extended to include around 4,000 interviews with 10-15 year olds. This extension is covered in more detail in [section 9](#).

The BCS is primarily a **victimisation** survey, in which respondents are asked about the experiences of **property crimes** of the household (e.g. burglary) and **personal crimes** (e.g. theft from a person) which they themselves have experienced. Since the move to

¹ Previous British Crime Surveys were carried out in 1982, 1984, 1988, 1992, 1994, 1996, 1998 and 2000.

continuous interviewing in 2001 the reference period for all interviews has related to the last 12 months before the date of interview. Although there have been changes to the design of the survey over time, the wording of the questions that are asked to elicit victimisation experiences, have been held constant throughout the life of the survey.

Respondents are asked directly about their experience of crime, irrespective of whether or not they reported these incidents to the police. As such the BCS provides a record of peoples' experiences of crime which is unaffected by variations in reporting behaviour of victims or variations in police practices of recording crime. The BCS and police recorded figures should be seen as a complementary series, which together provide a better picture of crime than could be obtained from either series alone. Crime statistics (including both the BCS and recorded crime statistics) have recently been subject to a number of reviews;

- Crime Statistics: User Perspectives, Report No.30 Statistics Commission, September 2006
- Crime Statistics: An independent review, November 2006
- Overcoming Barriers to Trust in Crime Statistics: England and Wales, UK Statistics Authority, May 2010

The extension of the survey to include 10-15 year olds was recommended by the two reviews conducted in 2006. The Home Office have also implemented changes in the way the results are released to the press as a result of these reviews.

The scope of the BCS goes well beyond the counting of criminal incidents, although it is for this estimate that it has become established as a definitive source of information. In order to classify incidents, the BCS collects extensive information about the victims of crime, the circumstances in which incidents occur and the behaviour of offenders in committing crimes. In this way, the survey provides information to inform crime reduction measures and to gauge their effectiveness.

As well as providing estimates of victimisation, another use of the survey has been to collect information to measure a number of performance targets, both at national and at individual police force level. The BCS was used as part of the Police Performance and

Assessment Framework (PPAF) to measure individual forces progress against a number of Statutory Performance Indicators (SPIs)².

1.2 Outputs from the BCS

The data arising from the BCS are mainly reported by the Home Office's Statistics Unit. These reports include:

- A full statistical bulletin based on BCS interviews carried out in the last financial year, which is published in the summer following the end of each financial year. This bulletin contains estimates from both the BCS and police-recorded crime figures. The latest of these reports covering the period 2009-10 was published in July 2010³, and can be found at:
 - <http://rds.homeoffice.gov.uk/rds/pdfs10/hosb1210.pdf>
- Shorter statistical updates produced on a quarterly basis, focusing specifically on victimisation rates and trend patterns
- Supplementary bulletins covering topics in the annual volume in more detail. Three supplementary volumes covering topics such as perceptions of anti-social behaviour, , Drug misuse, acquisitive crime and plastic card fraud have been published based on the 2008-09 BCS and can be found at:
 - <http://rds.homeoffice.gov.uk/rds/pdfs09/hosb1709.pdf>
 - <http://rds.homeoffice.gov.uk/rds/pdfs09/hosb1209.pdf>
 - <http://rds.homeoffice.gov.uk/rds/pdfs10/hosb0810.pdf>
- An annual bulletin covering drug misuse as reported on the BCS. The most recent bulletin for the period 2009-10 can be found at:
 - <http://rds.homeoffice.gov.uk/rds/pdfs10/hosb1310.pdf>

The above references are intended only to illustrate the types of reports and findings that are produced from the BCS. For more details on all RDS publications associated with the BCS see <http://www.homeoffice.gov.uk/rds/bcs1.html>

As well as published reports the BCS data is made available through the UK Data Archive at the University of Essex (<http://www.data-archive.ac.uk/>). Since considerable emphasis is given in the course of conducting the interview to assure respondents that

² For details on the PPAF see <http://police.homeoffice.gov.uk/performance-and-measurement/performance-assessment>

³ Flatley, J *et. al.* (eds.) Crime in England and Wales 2009-10 Home Office Statistical Bulletin 12/10

the information they provide will be held in confidence the data set does not identify the location of the sampled areas and this information is not released to the Home Office by TNS-BMRB.

The BCS is a complex study with data organised at different levels (households, individuals, and incidents) and it has numerous sub-samples that are asked specific questions. Accordingly considerable effort and expertise is required to analyse the data and to interpret it in a valid manner. Some of the analysis routines that play a key role in the published estimates are implemented after the data have been handed over to the Home Office, and are not documented in this report.

The Home Office produces a user guide for those interested in analysing BCS data which contain further detail on the content and structure of the data and guidance on analysis. This is also available from the UK Data Archive⁴.

1.3 Structure of the technical report

This report documents the technical aspects of the 2009-10 BCS carried out in England and Wales. The analysis in this report relates to the total sample that was issued in the financial year 2009-10, irrespective of when interviews actually took place. The distinction between issued sample and achieved sample is explained in more detail in [section 4.3](#) of the report.

The sample design is set out in [Chapter 2](#). Data collection is the major task for the organisation commissioned to conduct the BCS and forms the central part of this report. [Chapter 3](#) covers the content and development of the questionnaire, while [Chapter 4](#) examines the fieldwork. [Chapter 5](#) gives details of the tasks that are involved in preparing the data for analysis, including the coding and offence classification and [Chapter 6](#) covers the preparation and delivery of the BCS data files. [Chapter 7](#) outlines the weighting required for analysis of the data. [Chapter 8](#) provides the results of some checks on the profile of the BCS achieved sample against estimates for the population that the BCS aims to represent. [Chapter 9](#) details the extension of the survey to under 16s.

⁴ For the most recent User Guide see <http://www.data-archive.ac.uk/doc/6066/mrdoc/pdf/6066userguide.pdf>

2. Sample Design

2.1 Introduction

The sample design of the British Crime Survey remained largely unchanged between 2008-09 and 2009-10. A revised sample design was introduced to the survey in 2008-09 and full details of the rationale for the revised design and the design itself are included in the 2008-09 technical report. The key features of the 2009-10 design were as follows:

- A sample size of approximately 46,000 interviews per year with adults aged 16+ living in private residential households in England and Wales;
- A minimum of around 1,000 interviews per year in each of the 42 Police Force Areas⁵. This required a degree of over sampling in less populous Police Force Areas;
- A partially clustered design with different levels of clustering being used in different population density strata in an effort to reduce PSU-level cluster effects;
- Fieldwork was conducted on a continuous basis with sample being allocated to provide nationally representative estimates on a quarterly basis; and
- Sample was front loaded within each quarter to reduce the spill over of cases which are issued in one quarter but are interviewed in the next.

2.2 Target issued and achieved sample in Police Force Areas

A core requirement of the sample design was to achieve around a minimum of 1,000 interviews in each Police Force Area. The design which meets this requirement at minimum cost is one which delivers an equal sample of 1,000 interviews in each of the 42 Police Force Areas, giving an overall national sample of 42,000 interviews per year. However, such a design would result in a large range of sampling fractions (and hence design weights) within PFAs, leading to a reduction in the precision of whole sample estimates. It was therefore decided to adopt a design that boosted the sample size in smaller PFAs but without reducing it in the larger Areas compared to what it had been on previous surveys.

⁵ For sampling purposes the City of London Police are combined with the Metropolitan Police

This broad approach to over sampling in less populous Police Force Areas is the same one that has been adopted on the BCS since 2004-05 when the survey increased in sample size from 37,000 to 46,000. In 2008-09 the process was made slightly more systematic by allocating issued sample to the larger Areas in proportion to their population and this approach was repeated in 2009-10. With this approach the overall design effect was calculated at 1.17 using the standard formula that ignores between strata differences in element variance⁶.

Full details about the extent of over sampling within each PFA is contained in the 2008-09 Technical Report. The actual number of interviews achieved and the response rate for each PFA in 2009-10 is shown in [Table 4.13](#).

2.3 A partially clustered sample

The partially clustered sample design involves different sampling plans for each of three population density strata in an effort to reduce PSU-level cluster effects. The sample plans are defined as follows:

- In the **most densely populated** areas of each PFA an unclustered sample of addresses is drawn (Stratum A);
- In areas of **medium population density** a two-stage design is employed, first sampling Medium Layer Super Output Areas (MSOAs) as the primary sampling units and then selecting 32 addresses within each PSU (Stratum B); and
- In areas of **low population density** a three-stage design is employed by first sampling Medium Layer Super Output Areas (MSOAs), then selecting 2 Lower Level Super Output Areas (LSOAs) within each sampled MSOA as the primary sampling units, and finally selecting 16 addresses within each PSU (Stratum C);

2.4 Sampling of addresses

A different procedure for sampling addresses was adopted in each density stratum. All addresses were selected from the small-user Postcode Address File (PAF).

⁶ Formula is $(\sum n_h W_h)^2 / \sum n_h W_h^2$, where n_h = target sample size in PFA h and W_h = number of PAF delivery points in PFA h as a proportion of the total number of PAF delivery points in England and Wales

Sampling of addresses in the unclustered Stratum A

Within each PFA all the addresses allocated to unclustered stratum A were sorted using the ONS reference for the associated LSOA. Addresses were then sampled systematically using the PFA-level sampling fraction and a random start.

A geographic software system was then used to 'batch' together sampled addresses into efficient fieldwork assignments. In doing this certain parameters were set concerning the maximum geographic diameter of a batch area and the number of addresses per batch. The aim was to achieve assignments of a manageable geographical size that contained as close as possible to 32 addresses.

Census-derived and other government data were added to each batch using a weighted average of component LSOAs. This is best illustrated using an example. If a batch contained 8 addresses from LSOA 1, 16 from LSOA 2, and 9 addresses from LSOA 3 and the crime index values for each LSOA were 20, 30, and 40 respectively, the batch level crime index value would be:

$$(20*(8/33)) + (30*(16/33)) + ((40*(9/33)) \text{ or } 30.3$$

These batch-level data allowed a representative sample of batches to be allocated to each fieldwork quarter using standard stratification methods.

Sampling addresses in mid-clustered Stratum B

Before sampling, MSOAs in mid-cluster stratum B areas were stratified in the master database to ensure a representative sample. In England, mid-cluster MSOAs in each PFA were sorted by the crime and disorder deprivation index and split into three equal-sized sub-strata. In Wales, mid-cluster MSOAs in each PFA were sorted by population density and split into three equal-sized sub-strata.

These variables were selected after an analysis of BCS data from 2003-06 (see section 2.7 for further details).

MSOAs were sampled with a probability proportionate to the number of PAF delivery points⁷, using a systematic method and a random start.

32 addresses were selected from each sampled MSOA. Addresses were sorted by postcode before a systematic 1 in *n* sample was drawn with a random start.

⁷ In England and Wales, one delivery point equals one address in 97% of cases.

Sampling addresses in tightly clustered Stratum C

A sample of MSOAs was drawn in each tight-clustered stratum C as described for the mid-clustered strata. However, instead of a sample of addresses being drawn within each sampled MSOA, a pair of LSOAs was first selected.

Within each sampled MSOA, the component LSOAs were sorted using the ONS reference number. Two LSOAs were sampled in each MSOA with a probability proportionate to the number of PAF delivery points, using a systematic method and a random start.

Sixteen addresses were selected from each sampled LSOA. Addresses were sorted by postcode before a systematic 1 in n sample was drawn with a random start.

2.5 Stratification

The selection of PSU-level stratification variables was refined after an analysis of BCS data from April 2003 through to March 2006.

The same stratification was used in 2009-10 as was applied in 2008-09. This required the sample to be stratified by:

- PFA (level 1)
- Density cluster type (level 2)
- Three-band version of the 'crime and disorder' deprivation index (level 3) – England only
- Three-band version of population density (level 3) – Wales only

2.6 Allocation of sample to fieldwork quarters and months

Primary sampling units (mid and tight clustered strata) and fieldwork batches (unclustered strata) were systematically allocated to each fieldwork quarter to ensure that each quarter was a representative sample of the whole.

The sampled PSUs/batches in each cluster stratum were sorted using their original stratification values and tagged with a 'fieldwork quarter' label via the 'snaked' allocation system: Q1-Q2-Q3-Q4-Q4-Q3-Q2-Q1-Q1-Q2 etc. but with a random start (e.g. 'Q3').

A similar system was used to allocate sampled PSUs/batches to a specific issue month within the relevant quarter. However, rather than allocating PSUs/batches equally between months within each quarter the sample was slightly frontloaded within each quarter. This was done to try and increase the proportion of interviews that are actually carried out during the quarter of issue, rather than being carried out in the quarter after

issue. Thus, approximately 40% of the sample was allocated to month 1 of each quarter, 35% to month 2 and 25% to month 3.

2.7 Sampling of individuals

At each sampled address, interviewers were asked to randomly sample one dwelling unit in those rare cases where more than one is associated with a single address. Once the dwelling unit was selected, interviewers were asked to randomly sample one normally resident⁸ individual aged 16+. This was done by listing all eligible people in the household in alphabetical order of first name and then selecting one for interview by a random (Kish grid based) approach. Once the selection of an individual had been made no substitutes were permitted.

2.8 Extension to include 10-15 year olds

The 2009-10 survey was the first full year of the survey to include interviews with 10-15 year olds. The aim was to conduct around 4,000 interviews with young people aged between 10 and 15 years old. Screening for 10-15 year olds was conducted at core sample addresses and where possible a second interview was conducted at the address with a 10 to 15 year old. Full details of the 10 to 15 year old extension are included in [chapter 9](#).

⁸ An individual is 'normally resident' if this is his/her only residence or he/she spent more than six of the last twelve months living at this address.

3. Questionnaire Content and Development

3.1 Structure and coverage of the questionnaire

The BCS questionnaire for the adult survey has a complex structure, consisting of a set of core modules asked of the whole sample, a set of modules asked only of different sub-samples, and self-completion modules asked of all 16-59 year olds. Within some modules there is often further filtering so that some questions are only asked of even smaller sub-samples. The precise modules asked on the survey vary from year to year.

The 2009-10 BCS questionnaire consisted of the following sections:

- Household Grid
- Perceptions of crime
- Screener questionnaire
- Victimization Modules for incidents identified at the screeners (up to a maximum of six)
- Performance of the Criminal Justice System
- Mobile phone, second home and bicycle crime
- Experiences of the police (Module A)
- Attitudes to the Criminal Justice System (Module B)
- Crime prevention and security (Module C)
- Ad-hoc crime topics (Module D)
- Plastic card fraud
- Identity fraud
- Anti social behaviour
- Demographics and media
- Self-completion module on drug use and drinking
- Self-completion module on inter-personal violence

The basic structure of the core questionnaire is shown in Figure 3.1, while the sub-set of respondents who were asked each module of the questionnaire is shown in Table 3.1. The complete questionnaire is documented in Appendix D of Volume 2. In this chapter a brief description of each section or module of the questionnaire is outlined.

Figure 3.1 Flow Diagram of the 2009-10 BCS Core Questionnaire

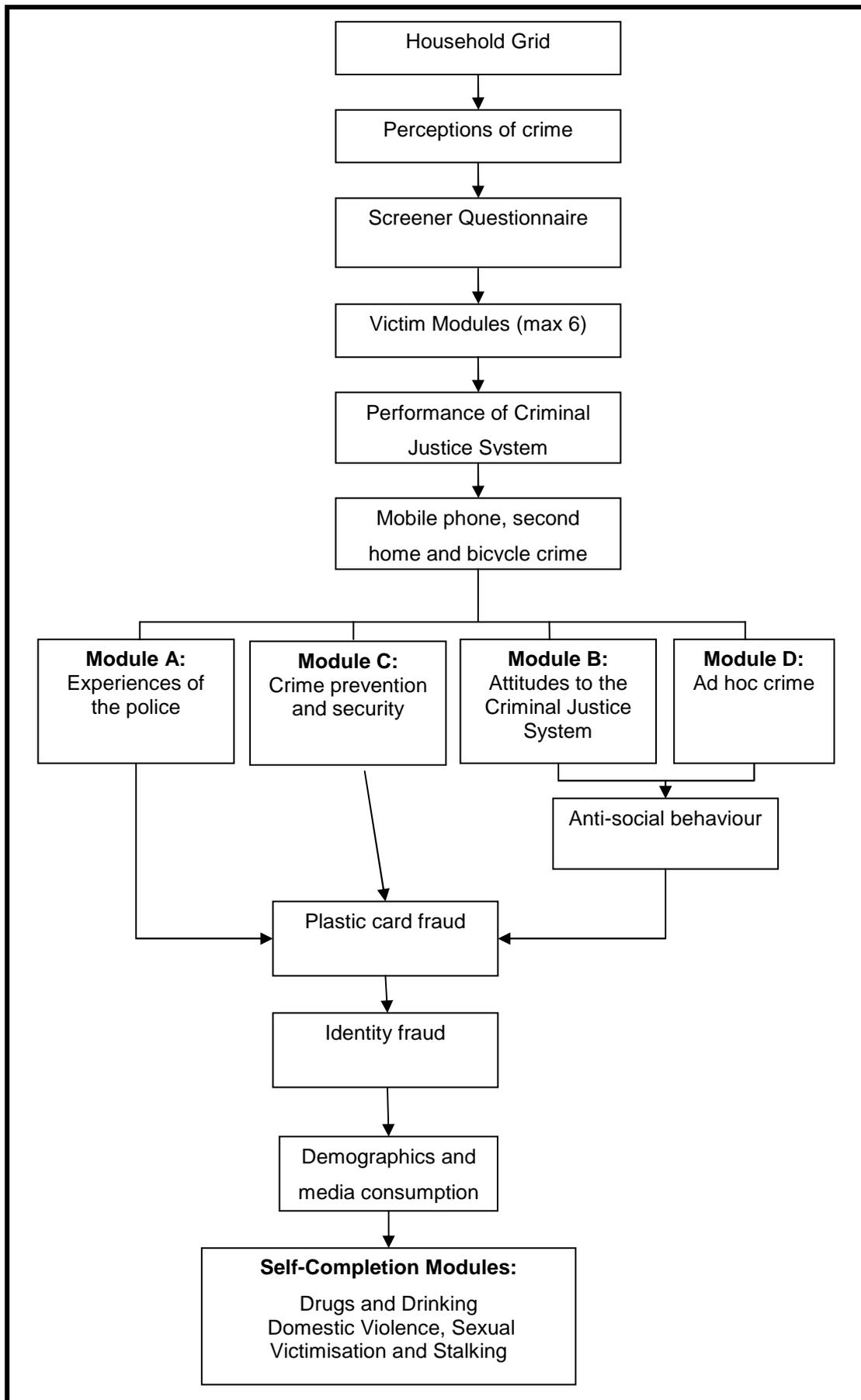


Table 3.1 Modules of the 2009-10 BCS questionnaire and sub-set of respondents who were asked each module

Questionnaire module	Core sample
Household box	All
Perceptions of crime	All
Screening questionnaire	All
Victimisation Modules	All victims
Mobile phone and second home crime	All
Performance of the Criminal Justice System	All
Module A	Random 25%
Module B	Random 25%
Module C	Random 25%
Module D	Random 25%
Anti social behaviour	Random 50%
Plastic card fraud	All
Identity fraud	All
Road safety and traffic ¹	All
Demographics and media consumption	All
Drugs and Drinking	All aged 16-59 ²
Inter-Personal Violence	All aged 16-59 ²

¹ Introduced in October 2009

3.1.1 Household Grid

Basic socio-demographic details (age, sex, marital status, etc.) were collected in the Household Grid for every adult in the household. Additionally, demographic details of all children under 16 years including their relationship with the respondent were collected.

The Household Grid was also used to establish the Household Reference Person⁹. Household Reference Person (HRP) is the standard classification used on all government surveys and is based on the following criteria:

- The HRP is the member of the household in whose name the accommodation is owned or rented, or is otherwise responsible for the accommodation. In households with a sole householder that person is the HRP.

⁹ Prior to 2001 all previous surveys collected details of the Head of Household.

- In households with joint householders the person with the highest income is taken as the HRP.
- If both householders have exactly the same income, the older is taken as the HRP.

3.1.2 Perceptions of crime

The Household Grid was followed by a series of attitudinal questions which asked respondents their perceptions about particular aspects of crime and anti-social behaviour. This module of questions included both long-standing questions, as well as questions first introduced on the 2008-09 survey.

Long-standing topics covered in this module included:

- How long respondents had lived in their local area;
- What respondents felt were the main causes of crime (Module B respondents only);
- Respondents' perceptions of changing crime levels for different types of crimes for the country as a whole and in their local area;
- How much crime and fear of crime affected respondents quality of life (Module D respondents only);
- How safe respondents felt when walking in their local area and when at home;
- How worried they were about being the victim of particular types of crime (Module C respondents only);
- How respondents thought crime rates in their local area had changed over time (Module C respondents only);
- How much of a problem they perceived particular aspects of anti-social behaviour to be;
- How often their home was left unoccupied and how often they went out; and
- How often they visited a pub or bar

3.1.3 Screener questions

Following the questions on perceptions of crime, all respondents were asked whether they had experienced certain types of crimes or incidents within a specified reference period, namely the last 12 months from the date of interview. To try and encourage respondents to recall events accurately, a life event calendar was offered to all respondents to act as a visual prompt when answering the screener questions (see section 3.2).

Depending upon individual circumstances a maximum of 25 screener questions were asked relating to particular types of crime. These can be grouped into four main categories:

- All respondents who lived in households with a vehicle or bicycle were asked about experience of vehicle-related crimes (e.g. theft of vehicle, theft from vehicle, damage to vehicle, bicycle theft);
- All respondents who had moved in the reference period were asked about experience of property-related crimes in their previous residence(s) (e.g. whether anything was stolen, whether the property was broken into, whether any property was damaged);
- All respondents were asked about experience of property-related crimes in their current residence; and
- All respondents were asked about experience of personal crimes (e.g. whether any personal property was stolen, whether any personal property was damaged, whether they had been a victim of force or violence or threats).

The wording of the screener questions has been kept consistent since the BCS began to ensure comparability across surveys. They are designed to ensure that all incidents of crime within the scope of the BCS, including relatively minor ones, are mentioned. The screener questions deliberately avoid using terms such as 'burglary', 'robbery', or 'assault', all of which have a precise definition that many respondents might not be expected to know.

The questions are also designed to ensure that the respondent does not mention the same incident more than once. At the end of the screener questions, the interviewer is shown a list of all incidents recorded and is asked to check with the respondent that all incidents have been recorded and nothing has been counted twice. If this is not the case, the respondent has an opportunity to correct the information before proceeding.

Within the screener questions a crucial distinction exists between household incidents and personal incidents.

All vehicle-related and property-related crimes are considered to be household incidents, and respondents are asked about whether anyone currently residing in the household has experienced any incidents within the reference period. A typical example of a household incident is criminal damage to a car. It is assumed that the respondent will be able to recall these incidents and provide information even in cases

where he/she was not the owner or user of the car. For respondents who have moved within the last 12 months, questions on property-related crimes are asked both in relation to the property they are now living in, as well as other places they have lived in the last 12 months.

Personal incidents refer to all crimes against the individual and only relate to things that have happened to the respondent personally, but not to other people in the household. An example of a personal incident would be a personal assault. An assault against other household members would not be recorded, unless the respondent was also assaulted in the course of the incident. In such cases, the offence would be coded according to the crime experienced by the respondent (which may not be the same as the experience of another household member).

3.1.4 Victimisation Modules

All incidents identified at the screener questions are followed through in more detail in the Victimisation Module. Incidents are covered in a specific priority order as explained below, which has been kept consistent since the start of the BCS.

Identification and ordering of incidents for Victimisation Modules

In 2009-10, 77% of core sample respondents did not report any victimisation over the reference period, meaning that no Victimisation Modules had to be completed as part of the interview. This is very similar to the proportion of respondents who did not report any victimisation in the 2008-09 survey (76%).

Where a respondent had experienced one or more incidents in the reference period, the CAPI programme automatically identified the order in which the Victimisation Modules were asked. This meant that the interviewer had no discretion about the selection or order of the modules¹⁰.

If six or fewer incidents were identified at the screener questions then a Victim Module was completed for all of the incidents reported. The priority ordering used by the computer was as follows:

¹⁰ In the case of the incidents of sexual victimisation or domestic violence, the interviewer had an option to suspend the Victimisation Module, as this might embarrass or endanger the respondent in some situations. The interviewer would then attempt to arrange a revisit at a time that would be more convenient (in particular when other household members would not be present).

- According to the type of crime. Victimization Modules were asked in reverse order to the screener questions. Broadly speaking this means that all personal incidents were asked before property-related incidents, which were asked before vehicle-related incidents.
- Chronologically within each type of crime. If a respondent reported more than one incident of the same type of crime, Victim Modules were asked about the most recent incident first and worked backwards chronologically.
- The first three Victimization Modules were long modules, which contain all the detailed questions relating to each incident. The second three Victim Modules were short modules, a cut down version of the questions that are much quicker to complete.

If the respondent had experienced more than six incidents in the reference period, only six Victimization Modules were asked using the above priority ordering. The priority ordering means that the survey does not collect details or only collects limited details (through the short Victim Module) for the crimes or incidents that tend to be more common (e.g. criminal damage to vehicles).

In the 2009-10 survey, a total of 14,950 Victim Modules were completed on the core sample and 23.2% of all respondents reported at least one incident (see Table 3.2).

Table 3.2 shows that 17% of all core respondents completed one Victimization Module, while only 1% of all respondents completed four or more modules. Among respondents who reported at least one crime, seven in ten (72%) had experienced only one crime in the reference period and so had completed a single Victimization Module. Only 4% of respondents who had been the victim of crime completed four or more Victim Modules.

Table 3.2 Number of respondents who completed Victimization Modules by sample type, 2009-10 BCS

	Core sample		
	N	% of all respondents	% of victims
Non victims	34,715	76.8	
Victims ¹	10,474	23.2	
No. of Victim Modules ²			
1	7,581	16.8	72.4
2	1,954	4.3	18.7
3	556	1.2	5.3
4	213	0.5	2.0
5	79	0.2	0.8
6	91	0.2	0.9
<i>Bases:</i>		45,189	10,474
<small>1 Victims refers to the number of respondents who completed at least one Victimization Module 2 The number of Victimization Modules is shown both as a percentage of all respondents who were victims of crime and as a percentage of all respondents.</small>			

Defining a series of incidents

Most incidents reported represent one-off crimes or single incidents. However, in a minority of cases a respondent may have been victimised a number of times in succession. At each screener question where a respondent reported an incident, they were asked how many incidents of the given type had occurred during the reference period. If more than one incident had been reported, the respondent was asked whether they thought that these incidents represented a 'series' or not. A series was defined as "*the same thing, done under the same circumstances and probably by the same people*". Where this was the case, only one Victimization Module was completed in relation to the most recent incident in the series.

There are two practical advantages to this approach of only asking about the most recent incident where a series of similar incidents has occurred. First, since many (although not all) incidents classified as a series tend to be petty or minor incidents (e.g. vandalism) it avoids the need to ask the same questions to a respondent several times over. Secondly, it avoids 'using up' the limit of six Victimization Modules on incidents which tend to be less serious.

In 2009-10, 82% of all Victimisation Modules related to single incidents and 18% related to a series of incidents. This split between single and series incidents was broadly the same as previous surveys.

In the rare cases where a respondent has experienced a mixture of single incidents and a series of incidents the interview program has a complex routine which handles the sequence of individual and series incidents and allows the priority ordering of the Victimisation Modules to be decided.

In terms of estimating the victimisation rates, series incidents receive a weight corresponding to the number of incidents up to a maximum of five (see section 7).

Content of Victimisation Module

The Victimisation Module is the key to the estimate of victimisation and collects three vital bits of information:

- The exact month(s) in which the incident or series of incidents occurred. In a few cases, respondents may have reported an incident, which later turned out to have been outside the reference period. In such cases, the Victimisation Module was simply by-passed by the computer. If respondents were unsure about the exact month in which something happened, they were asked to narrow it down to a specific quarter. For incidents that were part of a series, respondents were asked how many incidents occurred in each quarter and the month in which the most recent incident had occurred.
- In the questionnaire, program reference dates were automatically calculated based on the date of interview and appropriate text substitution was used to ensure that the questions always referred to the correct reference period. Because the 12 month reference period changed throughout the fieldwork year, this meant that some date-related questions in the Victimisation Module had different text each month to reflect this changing reference period. Details of these questions and the appropriate reference periods used for each month of the 2009-10 survey can be found in Appendix F of Volume 2.
- An open-ended description of the incident where the respondent describes exactly what happened in their own words. The open-ended description is vital to the accurate coding of offences that takes place back in the office. Short, ambiguous or inconsistent descriptions can often make offence coding difficult.

At the end of each Victimisation Module, the original open-ended description that the interviewer had entered at the start of the Victimisation Module is re-capped, along with the answers to some of the key pre-coded questions. By presenting this information on a single screen, interviewers have the chance to confirm with respondents that the information was correct and consistent. If the respondent and/or interviewer wish to add or clarify any information they then have the opportunity to do this.

- A series of key questions used to establish important characteristics about the incident. Examples of the sort of information collected includes where and when the incident took place; whether there was a racial element to the incident; whether anything was stolen or damaged and, if so, what; the costs of things stolen or damaged; whether force or violence was used and, if so, the nature of the force used and any injuries sustained; and whether the police were informed or not.

The key questions within the Victimisation Module have remained largely unchanged from previous years of the survey to ensure comparability over time.

3.1.5 Performance of the Criminal Justice System

All respondents were asked a number of questions about the performance of both the Criminal Justice System as a whole, as well as about the individual agencies that make up the CJS.

The first set of questions in this module relate to respondents' perceptions about the effectiveness and fairness of the CJS. Individual questions relating to the police, the courts, the CPS, the probation service and the prison service were asked, as well as questions about the CJS as a whole. These questions were added to the survey in October 2007 after being extensively tested¹¹.

The second set of questions is about confidence in the local police. As well as a general question about perceptions of how good a job the local police are doing, there are also questions related to specific aspects of local policing.

¹¹ Maxwell C. *et. al.* (2008) Fairness and effectiveness in the Criminal Justice System: development of questions for the BCS at <http://www.homeoffice.gov.uk/rds/pdfs08/doqbcsc.pdf>

Finally, the module includes a number of questions related to respondents' confidence in the different local agencies involved in tackling crime and anti-social behaviour. These questions are used to measure progress at a national level towards PSA Delivery Agreement 23 which aims to make communities safer. The questions are also used to measure the performance of individual police forces in England and Wales and their progress towards a 2012 target that 60% of the public are confident that their local police are addressing the crime and anti-social behaviour issues that matter to them.

3.1.6 Mobile phone, second home and bicycle crime

Although mobile phones stolen from the respondent should be identified in the Victimization Module, thefts from other members of the household are not covered. Consequently, in this module all respondents were asked who in the household (if anyone) used a mobile phone, whether anyone in the household had had a mobile phone stolen in the last 12 months and, if so, who the phone had been stolen from. Respondents were asked to include incidents where mobile phones stolen had been stolen from children in the household. A similar set of questions were added to the 2009-10 survey about bicycle theft from any members of the household. These questions were added to enable cross checks across the adult and 10 to 15 year old data to identify avoid any double counting of incidents.

The survey does not pick up crimes that affect second homes since respondents are only interviewed at their main address, and are asked about their main property. The module includes questions about whether respondents owned a second home in England and Wales.

For respondents who did have a second property (or properties) there were then a number of follow-up questions asking them whether this property has been broken into or suffered any form of criminal damage in the last 12 months.

Data from these questions should be analysed using the household weights to generate victimisation rates per household. These are discussed in greater detail in [Chapter 7](#).

3.1.7 Part-sample modules (A-D)

Respondents were randomly allocated to one of four modules (see section 3.5. for how this was done). Core sample respondents were allocated equally to each module, meaning that approximately 11,500 respondents were asked each module.

Module A: Experiences of the police

Module A included topics such as:

- whether or not respondents are serving police officers or had any contact with the police;
- where they obtain details about their local police force
- whether or not they had been stopped by the police either in a vehicle or on foot;
- if so, the reason given for being stopped and the nature of the contact;
- whether respondents had made a complaint about the police and if so how they felt their complaint had been dealt with; and
- vicarious contact with the police (added in 2009-10 survey)

In 2009-10 a number of extra questions were added to this module about the policing pledge, neighbourhood policing and awareness of local crime maps.

The questions covered:

- Whether respondents had noticed any change in how often police or police community support officers are seen on patrol in the local area.
- Whether respondents know how to contact the local police about policing, crime or anti-social behaviour issues
- Awareness of the neighbourhood policing team
- Whether they have heard about anyone who has had a bad experience with the police and where they heard it from.
- Whether seen, read or heard any information about the local police and if so where this was seen
- Whether heard of local crime maps and whether looked or used the maps

Module B: Attitudes to the Criminal Justice System

This module included questions that had mainly been asked in previous years, although there were some newer questions about the types of sentences that respondents thought appropriate for different types of offenders under particular circumstances. Topics covered in this module included:

- perceived leniency or toughness of the Criminal Justice System
- opinions as to what Criminal Justice System should do to improve confidence;
- attitudes to the type of sentence appropriate for different types of offenders under particular circumstances;
- attitude to sentencing policy, including what respondents thought sentences should be for particular crimes and what they thought they actually were;
- awareness and perceived effectiveness of Community Payback (previously known as community service);
- experiences of the Criminal Justice system; whether respondents have worked for CJS, have been arrested, been a juror, a victim in a case, etc; and
- attitudes to aspects of the Youth Justice System

Module C: Crime prevention and security

Topics covered in this module vary from year to year. In 2009-10 the main focus was on home security measures. Other topics covered in this module included.

- Neighbourhood watch
- personal security measures
- vehicle security measures

Questions on home security included some questions on security measures fitted to the home (e.g. burglar alarms, window locks, type of door locks, etc.) and action taken when a stranger comes to the door.

Module D: Ad hoc crime

This module was broadly similar to previous surveys and contained a wide variety of questions. These included:

- worry about gun crime and terrorism; and
- concerns about being the victim of certain crimes;
- social cohesion – the likelihood of people in the neighbourhood taking action if they saw crime or anti-social behaviour taking place

3.1.8 Plastic card fraud

This module is intended to provide a measure of the extent of plastic card fraud. This type of crime is not currently included in incidents covered in the Victimisation modules

(though the physical theft of any plastic card would be covered). The module first appeared on the survey in August 2007.

The topics covered in the module included:

- whether respondent has had a plastic card used without their permission;
- whether respondent has had money taken from a bank or building society account without their permission and the amount stolen;
- reporting of plastic card fraud
- measures taken to try and prevent card fraud

3.1.9 Identity Fraud

In a departure from previous years, questions relating to identity fraud were included in a separate module. It was deemed important to cover issues relating to identity fraud as a stand-alone issue and in more depth.

The topics included:

- whether respondent has had their personal details used in any way that indicated identity fraud;
- circumstances under which identity fraud occurred;
- who identity fraud was reported to
- how respondents believe their personal details were acquired

3.1.10 Anti-social behaviour module

This module includes some general questions on anti-social behaviour. These related to how effective respondents felt the authorities were in tackling anti-social behaviour, how informed respondents felt about what was being done locally to tackle problems.

Between April 2009 and October 2009 there were some questions related to three specific types of anti-social behaviour:

- teenagers hanging around;
- people using or dealing drugs; and
- people being drunk or rowdy in public

The main aim of these questions is to try and get a better understanding of the relationship between respondents' perceptions of anti-social behaviour in their local area and their actual experiences.

3.1.11 DfT Road safety and traffic module

From October 2009, for 12 months only, a module was added to the survey including questions about experiences of road accidents and problems with speeding traffic in the area.

Questions included:

- Whether been in a road accident in the last three years and whether injured in any road accident in the last three years and how many times this had happened
- How respondent was travelling at the time of the accident (in the car, as a cyclist, pedestrian, motorcyclist etc)
- Whether the police attended the incident and whether the incident was reported to the police.
- Why speeding traffic is a problem in the area
- What types of road speeding traffic is a problem on and the sorts of problems caused by it
- How much the respondent is affected by speeding traffic
- How many miles the respondent has driven in the last 12 months

3.1.12 Demographics and media consumption

This section collected additional information on the respondent and the Household Reference Person (where this was not the same as the respondent). Questions included:

- health condition;
- employment details;¹²
- educational attainment and qualifications;
- nationality, country of birth and religion (of respondent and HRP);
- ethnicity
- housing tenure; and
- household income

This section also covered media consumption habits.

Questions asked included:

- daily and preferred readership;

¹² Where the respondent was not the Household Reference person occupation details were also collected about the HRP

- amount of TV watched;
- amount of local and national news watched on TV¹³; and
- use of the internet and type of sites visited¹⁴

3.1.13 Self – completion modules

The self-completion modules are asked of respondents aged 16 to 59 years of age. The self completion modules are all presented as computer assisted self-completion (CASI) modules to ensure respondent confidentiality in answering these questions. The respondent was asked to follow the instructions on the screen of the laptop and enter their answers appropriately. Practice questions were included before the start of the self-completion module to give the interviewer an opportunity to show the respondent the different functions of the computer. If the respondent was unable or unwilling to complete the modules using the computer the interviewer could administer the self-completion. Where interviewers administered the self-completion, respondents were only asked the modules on drug use and drinking. They were not asked the module on domestic violence, sexual assault and stalking. Interviewer assistance and the presence of others while completing these modules was recorded by the interviewer (see [Chapter 4](#)).

Self-completion module: Illicit drug usage

The module covers a total of 17 drugs plus 3 more general questions to capture use of any other substances. The drugs included are:

1. Amphetamines
2. Methamphetamine
3. Cannabis
4. Skunk
5. Cocaine powder
6. Crack cocaine
7. Ecstasy
8. Heroin
9. LSD/Acid
10. Magic Mushrooms
11. Methadone
12. Semeron

¹³ These questions were removed in October 2009

¹⁴ These questions were removed in October 2009

13. Tranquilizers
14. Amyl Nitrate
15. Anabolic steroids
16. Glues, solvents, gas or aerosols
17. Ketamine
18. Unprescribed unknown Pills or powders
19. Smoked any substance (excluding tobacco)
20. Any other drug

The list of drugs included a drug that did not exist to attempt to identify instances of over reporting.

Questions included:

- whether ever taken illegal drugs;
- whether taken illegal drugs in last 12 months;
- whether taken illegal drugs in last month; and
- frequency of drug use
- frequency of drinking alcohol and how often felt drunk in the last 12 months

In October 2009 questions were added to the survey to record whether the respondent had taken legal highs in the last 12 months.

Self-completion module: Domestic violence, sexual victimisation and stalking module

The module was largely based on the module first developed in 2001 (and modified in 2004-05) to measure prevalence of domestic violence, sexual victimisation, and stalking.

The 2009-10 questions on inter-personal violence covered the following topics:

- experience of domestic violence by either a partner or by an other family member since age 16 and in the last 12 months;
- experience of less serious sexual assault since age 16 and in the last 12 months;
- experience of serious sexual assault since age 16 and in the last 12 months; and
- experience of stalking since age 16 and in the last 12 months

In 2009-10, those who had been subjected to serious sexual assault since the age of 16 were asked supplementary questions about the nature of the sexual assault. The questions covered:

- frequency of incidents;
- whether the police came to know or not;
- whether drugs or alcohol were involved;
- whether respondent suffered any injuries or sought any medical help; and
- whether respondent had to take any time off work

The module also includes a question on the respondent's sexual orientation.

3.2 Life event calendar

To aid respondent recall the BCS makes use of a life event calendar.

Such a calendar works by trying to place events or incidents in some sort of meaningful context for each respondent by building up a picture of events that have happened to them in the last year (e.g. birthdays, anniversaries, holidays, starting a new job, etc.) which are memorable to the respondent. Additionally, national dates such as Christmas, Easter, or Bank Holidays can be put on the calendar as common reference points. Further details about the thinking behind the life event calendar and its development can be found in the 2001 BCS Technical Report.

In relation to the BCS, the life event calendar can be used for two purposes:

- First, to provide respondents with a visual aid throughout the screener questions; and
- Second, to help respondents who were having difficulty recalling in which particular month an incident may have occurred.

Appendix E in Volume 2 has an example of the calendar used on the 2009-10 survey.

3.3 Questionnaire development

Since most of the questions on the 2009-10 BCS had been included in previous years of the survey, it was decided to concentrate piloting efforts primarily on new questions. In 2009-10 piloting was conducted using cognitive testing in central locations.

A first stage of cognitive testing was carried out by researchers. Cognitive testing uses probing techniques to try and understand the thought processes that a respondent uses in answering a survey question. It is designed to see whether the respondent understands the question, or specific words and phrases contained within the question; what sort of information the respondent needs to retrieve in order to answer the question; and what decision processes the respondent uses in coming to an answer.

The piloting was conducted in two rounds and was carried out in central locations to maximise the efficiency of the process. Interviewers carried out in-street recruitment according to broad quotas in town centres, while researchers carried out the cognitive interviewing using paper questionnaires. With several researchers able to carry out interviews at the same time, this method allowed about 20 cognitive interviews to be carried out in a single day. All researchers worked to the same probe guide and interviews were recorded.

Following the first round of testing the questions were revised and a second round of cognitive testing was carried out. The first round of piloting was conducted in Andover and Kingston and the second round was conducted in Brent Cross and Manchester. All respondents were given a £10 high street voucher to thank them for taking part in the pilot. In total 41 interviews were conducted across both rounds of piloting.

The main question areas covered in the 2009-10 piloting were as follows:

- Awareness of neighbourhood policing teams, whether they had seen information about the police, awareness and use of interactive crime maps
- Contact with the police, whether the police told them what action would be taken, how the police dealt with crimes reported and how respondents felt they were treated by the police
- Vicarious experiences of the police – whether respondents had heard about good or bad experiences of the police from other people
- Awareness of and attitudes towards asset recovery
- Awareness of and attitudes towards community payback
- Confidence that police and local authorities are tackling crime or anti-social behaviour committed by young people under 18
- Awareness of and use of online auction sites

The full pilot reports of the 2009-10 survey can be found in Appendix J of Volume 2

3.4 Final questionnaire and revisions

A paper questionnaire was produced from the Quanquest software that detailed the questions and their routing instructions as specified in the Quanquest code. This was translated into a Word document to provide a more user-friendly questionnaire.

Once all changes had been approved the questionnaire was thoroughly checked by TNS-BMRB researchers and Home Office research staff.

3.5 Allocation of sample within CAPI

In the 2009-10 survey the unique serial number entered by interviewers into the computer had to be capable of the following:

- to randomly allocate respondents to one of four part-sample modules (and within each module to further allocate respondents into a sub-sample)
- to distinguish between a core sample respondent and a young adult boost respondent

The unique serial number pre-printed on all core Address Contact Sheets and transferred by interviewers into the CAPI consisted of 6 digits. The first 4 digits (1000-9999) represented the area or sample point number and the last 2 digits (01-99) represented the address number. Additionally, the interviewers had to enter a screen number which denoted whether the interview was a core sample interview (screen number 0) or 10-15 year old interview (screen number 8). Various checks were incorporated into the questionnaire to minimise the chances of errors being made by interviewers when entering the serial and screen numbers.

Allocation of respondents to each part-sample module was done on the basis of the address number, using an algorithm based on division of the address number by 8 as shown in Table 3.3. The allocation to a particular Module was done automatically at the start of the interview by the CAPI programme when the interviewer entered the serial number.

Since each sample point contained 32 addresses the above algorithm ensured that within each sample point a similar number of issued addresses were randomly allocated to each follow-up module.

Table 3.3 Allocation of interviews to modules

Address Numbers	Remainder divided by 8	Allocated module
01/09/17/25	1	A1
02/10/18/26	2	B1
03/11/19/27	3	C1
04/12/20/28	4	D1
05/13/21/29	5	A2
06/14/22/30	6	B2
07/15/23/31	7	C2
08/16/24/32	8	D2

In the event this method of randomly allocating respondents to different sub-modules ensures that the process is strictly controlled and results in an even allocation across the year. Table 3.4 shows the actual proportion of respondents allocated in 2009-10 to the different sub-modules against the target.

Table 3.4 Achieved allocation of respondents to modules against target, 2009-10 BCS

Module	Target allocation	Achieved allocation
A1	12.5%	12.6%
B1	12.5%	12.5%
C1	12.5%	12.6%
D1	12.5%	12.4%
A2	12.5%	12.5%
B2	12.5%	12.5%
C2	12.5%	12.4%
D2	12.5%	12.5%

3.6 Features of Quancept used in the BCS

3.6.1 Don't Know and Refusal Keys

In the Quancept script, Don't Know and Refused are special codes. Rather than entering numeric codes for these options, interviewers enter DK and REF respectively. As with previous years of the survey, almost every question had a Don't Know and Refused option that the interviewer could use. However, at most questions they were hidden, and so did not appear on the screen as an explicit option to try and ensure that interviewers did not over use these options. In the paper questionnaire in Appendix D of Volume 2, Don't Know and Refused are only shown if they actually appeared as an option on the screen.

3.6.2 Different question types

The vast majority of questions were pre-coded, meaning that a list of answer categories appears on the laptop screen and the interviewers enter the appropriate numeric code. Questions were either single response (i.e. only one code can be entered) or multi-response (i.e. more than one code can be entered). In the latter case, entered answers are separated by spaces. In multi-response questions it is possible to allow a combination of either multi-response or single response options at the same

question. In the case of numeric questions, where an actual value is required, the interviewer simply types in the appropriate number.

Many pre-coded questions had an 'Other –specify' option, and if this option was selected by a respondent, the interviewer would simply type in the answer given. In all these questions, the answers were later examined by coders to see if the 'other' answer could be back coded into one of the original pre-coded options (see [section 5.2](#)).

In Quancept the standard keys that interviewers use to move forwards and backwards through the questionnaire are *Ctrl + Enter* and *Ctrl + Backspace* respectively. It was felt that these keystroke combinations might be awkward for respondents when completing the self-completion part of the questionnaire. Consequently, a modified version of the software was used for the BCS which allowed respondents to use single keystrokes (F2 for forward, F1 for backward) to complete the self-completion.

3.6.3 Logic and consistency checks

A number of logic and consistency checks were built into the Quancept script. These were of two types: hard checks and soft checks. Hard checks are ones where the interviewer is unable to move to the next question until the discrepancy or inconsistency has been resolved. Soft checks are ones where the interviewer is asked to confirm that the information entered at a specific question is correct but is able to pass on to the next question. An example of a hard check is to make sure that every household has someone coded as the Household Reference Person. Until the interviewer codes someone in the household as the HRP they cannot move forward. An example of a soft check is to check the value of stolen items that appear low (for example, a vehicle). In this case the interviewer will be prompted to check with the respondent whether the value entered is correct or not, and has the option either to change the original answer or leave it as it is.

A full list of all the logic and consistency checks in the 2009-10 script can be found in Appendix I of Volume 2.

3.6.4 Date calculation and text substitution

Text substitution and date calculations were used extensively throughout the questionnaire.

Text substitution is where alternative text is used in a question depending upon the series of answers given by a respondent to previous questions. In the paper questionnaire, square brackets are used to denote the existence of text substitution in a question.

Two main types of **date calculations** were used in the questionnaire:

- First, the precise reference period was calculated based on the date of interview and this was then substituted into the text of many questions. In all cases it was decided to calculate the date to the first of the month 12 months previous. Thus, for example, any interviews conducted in July 2009 would use the reference period “*since the first of July 2008*”. This means that in practice the 12 month reference period consisted of the last 12 full calendar months, plus the current month (i.e. slightly more than 12 months). This fact is taken into account when the victimisation rates are estimated.
- Second, some code frames consisted of particular time periods (e.g. months or quarters) which changed on a month by month basis. With these type of questions the Quancept script was programmed to allow the whole reference period covered by the questionnaire (that is, from April 2008 to June 2010 – a total of 27 months). However, interviewers only saw on screen the sub-set of codes that were appropriate to the correct reference period (i.e. 13 calendar months) for the month they were interviewing in.

Since some questions use these constantly rotating code frames based upon date of interview it is impossible to label these variables in any meaningful way in the SPSS data file. A list of these questions and the appropriate code frames that actually appeared on screen depending upon the month of interview can be found in Appendix F of Volume 2.

4. Fieldwork

This chapter documents all aspects of the data collection process, focusing on fieldwork procedures, the management of fieldwork across the survey year, quality control procedures and response rates achieved across the different samples.

4.1 Briefing of interviewers

All interviewers working on the 2009-10 survey attended one of two types of briefings during the year. Interviewers who had not previously carried out a BCS assignment were required to attend a full day face-to-face briefing before they could work on the survey. These briefings were held throughout 2009-10 as required. In total, 6 full briefings were held and 88 interviewers were briefed during the year.

New interviewers were also asked to attend a half-day briefing about six months or so after they had finished their first BCS assignment. This was intended to be a chance for new interviewers to seek clarification about any of the field procedures they were unsure about; to share experiences and good practice amongst each other; and generally to provide new interviewers with a supportive environment for developing their skills.

All briefings were presented by TNS-BMRB researchers and field staff working on the survey, and some were also attended by Home Office staff.

Each briefing covered the following topics:

- some background to the BCS and how the information is used by the Home Office;
- details about sampling and fieldwork procedures and advice on how to obtain high response rates;
- an introduction to the Address Contact Sheet and how to carry out the selection procedures;
- an introduction to the BCS questionnaire. The primary purpose of this part of the briefing was not to cover every single question in the survey but to cover the broad structure of the questionnaire and provide key pointers on how to collect accurate and comprehensive information from the screener questions and the Victimization Module. Additionally, this part of the briefing looked at how interviewers should approach the self-completion sections of the questionnaire.

- 10-15 sample: an explanation of the screening and selection procedures for 10 to 15 year olds; the field documents (leaflets, parental information cards, etc.); a discussion of the consent procedures to be used on the survey; and going through the questionnaire, including the use of CASI and audi-CASI.

In addition to this face-to-face briefing, before starting their BCS assignment interviewers were also required to read the written Interviewer Instructions and carry out at least two practice interviews based on particular scenarios provided in the Instructions.

It is normal practice on BCS to brief experienced BCS interviewers at least once a year, holding a half-day 'refresher' briefing. The last set of refresher briefings were held in August and September 2009.

These refresher briefings covered:

- Additional background information about the survey, including an update on the latest results published
- Update on response rate and levels of achievement across the country
- Recent changes to the questionnaire
- Data security
- Hints and tips on improving response among Under 16s

A total of 23 refresher briefings were held in 2009-10 and 338 interviewers attended.

4.2 Supervision and quality control

Several methods were used to ensure the quality and validity of the data collection operation.

A proportion of interviewers, particularly those less experienced, were accompanied in the field by supervisors. This included interviewers who were new to random probability sample surveys, who were accompanied on the first day of their BCS assignment by a supervisor. A total of 225 interviewers working on a BCS assignment were accompanied by a supervisor during 2009-10.

A proportion of addresses were re-contacted, to verify that the interviewer had contacted someone at the address and whether or not an interview had resulted. In

total, 6,087 addresses were re-contacted (13% of addresses where an interview was achieved) to verify that the interviewer had contacted someone and to determine whether an interview had resulted. Addresses for back checking were selected on the basis of TNS-BMRB's overall field quality procedures, whereby all interviewers have their work checked at least twice a year. A total of 419 separate BCS assignments were back checked during the year.

Validation was carried out mainly by telephone. Where no telephone number was available a short postal questionnaire was sent to the address to collect the same information.

4.3 Fieldwork dates and fieldwork management

During 2009-10 the survey was managed on a monthly basis. As mentioned in Section 2.8, it was decided to frontload the sample on a quarterly basis rather than issuing an even number of assignments each month. Thus, approximately 210 assignments were issued at the start of the first month of each quarter, 180 assignments were issued at the start of the second month, and 140 assignments were issued at the start of the third month. The aim of this approach was to try and get a balance between on the one hand reducing the proportion of interviews where sample was issued in one quarter, but the interview was conducted in the following quarter; but on the other hand maintaining a relatively even flow of interviews throughout the year.

Interviewers were encouraged to start their assignment as early as possible in the month to minimise the time between respondents receiving the advance letter and an interviewer calling. Interviewers had until the end of the calendar month to cover all the addresses in their assignment and report final outcomes.

Once all the issued addresses had been covered the Address Contact Sheets were returned to Head Office and a decision was taken about re-issuing non-productive outcomes. As a general rule all non-productive addresses (non-contacts, refusals, broken appointments, etc.) were re-issued unless there was a specific reason not to or it was considered not to be cost effective (e.g. only one or two addresses in an assignment). Once the first re-issue period had been completed a decision was taken about whether to re-issue addresses that were still non-productive for a second or third time.

In total across the year, 13,047 addresses were re-issued on the core sample, which represented 20% of the original sample. Of these 3,874 addresses were issued for a second time (6% of all addresses), and 685 (just over 1% of addresses) were issued for a third time. Just 33 addresses were issued a fourth time. Of all the addresses re-issued, 22% were converted into productive outcomes at some stage. Addresses where the original outcome had been a refusal were less likely to be converted than those that had been a non-contact or some other unproductive outcome (e.g. broken appointment, away, etc.). Overall, the impact of the re-issue process was to increase the response rate on the core sample from 70.1% after the initial issue to the final response rate of 75.6% (see [section 4.7.1](#)).

Because of this time lag between addresses being issued and interviews being achieved, the time period covered by the 2009-10 issued sample and the time period covered by the 2009-10 achieved sample are different. Although sample for the survey was issued between April 2009 and March 2010, the actual fieldwork dates over which interviews were achieved ran from April 2009 to June 2010. As already explained this means that for each quarter of the year not all interviews were actually achieved in the quarter of issue. In fact, approximately 83% of interviews were achieved in the same quarter as they were issued, with 17% of interviews falling into the next quarter. Not surprisingly, most of the interviews that fell into the following quarter were those issued in the last month of a quarter (i.e. June, September, December and March).

A change in survey procedure implemented in 2008-09 continued in 2009-10 in that the questionnaire used in the field was aligned to the survey year, rather than being aligned to the sample issue. Before 2008-09, the exact questionnaire used for any individual interview depended upon the period in which the sample was issued.

In 2009-10 all interviews carried out from the 1st April were therefore done with the new 2009-10 questionnaire, irrespective of the time period in which the sample was issued. The advantage of this is that the questionnaire is now in line with the way in which the data are reported. This was also the case in October when mid-year changes to the questionnaire were introduced.

Further details of how the quarterly data outputs relate to the issued and achieved sample can be found in [section 6.2](#).

4.4 Fieldwork procedures and documents

All assignments in the clustered part of the sample consisted of 32 addresses. As part of the process to batch up the unclustered part of the sample into manageable fieldwork assignments an attempt was made to make assignments of 32 addresses wherever possible. However, in practice this was not always possible and so assignment sizes did vary. In fact, 89% of assignments in 2009-10 consisted of 32 addresses, while 92% had between 30 and 34 addresses. The largest assignment consisted of 40 addresses, while the smallest assignment consisted of 6 addresses.

4.4.1 Advance letter and leaflet

All selected addresses were sent a letter from the Home Office in advance of an interviewer calling at the address. For addresses in Wales, a Welsh translation was provided on the reverse of the letter. This explained a little about the survey, why this particular address had been selected and telling the occupiers that an interviewer from TNS-BMRB would be calling in the next few weeks. The letter also provided a telephone number and an email address for people to contact to find out more about the survey, to make an appointment for an interviewer to call, or to opt out of the survey. Over the course of the whole year 1,082 people, representing around 2% of addresses issued, opted out of the survey by contacting either TNS-BMRB or the Home Office.

Included with the advance letter was a leaflet from the Home Office which provided people with some more details about the survey, including findings from the previous survey. The leaflet also tried to answer some questions that potential respondents might have such as issues relating to confidentiality.

Examples of the advance letters used can be found in Appendix A and a copy of the leaflet can be found in Appendix B of Volume 2.

4.4.2 Address Contact Sheets (ACS)

Interviewers were issued with an Address Contact Sheet (ACS) for each sampled address. This was the key document that allowed interviewers to carry out the different tasks that make up the BCS assignment and to record and manage their own calling strategies for each address.

The Address Contact Sheets are crucial documents to the management of the BCS, both at the level of the individual assignment and for the management of the survey overall. The primary functions of the Address Contact Sheet are as follows:

- To allow interviewers to record the days and times that they called at an address. Additionally, there was space for interviewers to record details or comments that may be useful should the address be re-issued to another interviewer.
- To provide a record of all the outcomes achieved at the address. The ACS allowed the outcome at each re-issue stage to be recorded separately, so that there was a complete record of outcomes for each address. Although these outcomes were recorded by interviewers on the paper ACS, they were also reported electronically to Head Office on a daily basis so that overall progress could be monitored and managed.
- To allow the interviewer to carry out any selection procedures where required. Where an interviewer found more than one dwelling unit at an address they had to carry out a procedure to randomly select one dwelling unit for interview. Similarly, where more than one eligible adult was found at an address, interviewers had to randomly select one person for interview. The ACS allowed interviewers to carry out these procedures and record the details for future reference or checking.
- To allow the interviewer to carry out the screening process for the 10 to 15 year olds survey the ACS had step by step instructions for interviewers about how to carry out these procedures and also allowed them to record the screening outcomes for every address. As with the final response outcomes, all screening outcomes were reported back to Head Office on a daily basis.
- To collect some basic information about the area and the selected address (e.g. type of property, condition of the property, whether it is in a Neighbourhood Watch area, etc.). This information was collected by interviewers based on their own observations and, as such, was highly subjective. Nevertheless, such information does tend to be highly associated with non-response and is also used by the Home Office as an area based disorder measure. This observational data was recorded by interviewers on the back page of the ACS. Interviewers returned this information by completing a short CAPI survey for each address as part of their end of day administration procedures. The data was added to the main data files at a later stage.

Examples of the Address Contact Sheets can be found in Appendix C of Volume 2.

4.5 Presence of others during the interview

During the interviewer briefing sessions emphasis was given about trying, wherever possible, to conduct the interview in private. This generally helps to make the interview run more smoothly, but it also might encourage some respondents to mention certain incidents or events, which they might be embarrassed or worried of talking about in front of others.

Privacy during the interview is a particular concern for respondents who have experienced domestic violence or sexual assault. Where respondents had experienced such incidents in the last 12 months, interviewers had the option of suspending the Victimisation Module (simply by skipping over it) if they felt it was inappropriate to continue with the questions because of the presence of others in the room. This procedure meant that the interviewer could complete the rest of the questionnaire, rather than having to abandon the whole interview. During 2009-10, a total of 18 Victimisation Modules were suspended by interviewers for this reason.

Although it is preferable for the interview to be conducted with no-one else present, there are also some situations where the presence of others might improve the accuracy of the information collected. This is particularly the case in incidents of vehicle crime or property crime, where the respondent may not have been personally present, reported the incident to the police, etc. Additionally, in many cases it is simply not be possible for the interview to be conducted without others present in the room.

4.5.1 Presence of others during the screener interview

The key point at which the presence of another person could affect the estimate of victimisation is during the initial set of screener questions. Therefore, at the end of these questions, the interviewer recorded whether anyone else was present. Table 4.1 shows whether or not anyone else was present in the room during the initial screener questionnaire, when respondents are giving details about their experiences of crime.

Table 4.1 Presence of others during the screener questionnaire, 2009-10 BCS

Core sample	
	%
No-one present	70
Children under 16	8
Spouse/partner	17
Other adult	7
<i>Base: All respondents</i>	<i>45,189</i>

In 2009-10, seven out of ten (70%) respondents were interviewed with no-one else other than the interviewer being present. Where someone else was present, the people most commonly there were the respondent's spouse or partner (17%) or their children (8%).

There was little difference between men and women as to whether they completed the interview with no-one else being present (72% of men and 69% of women).

Asian respondents, and in particular Asian women, were less likely than respondents from other ethnic groups to have done the screener questionnaire with no-one else present. Thus, 55% of Asian respondents completed the screener with no-one else present, while only 47% of female Asian respondents did so (compared with 63% of Asian men). Asian respondents were more likely than other respondents to have a spouse or partner present (23%), children present (17%), or another adult present (15%).

However, any patterns by age or ethnicity will also be influenced by household composition. Table 4.2 shows the information from the previous table with single person households identified separately.

Not surprisingly this shows that the vast majority of respondents interviewed in single person households were interviewed with no-one else present. The majority of respondents living in households with more than one person were also interviewed with no-one else present, although around four in ten respondents were interviewed with someone else present.

Table 4.2 Presence of others during the screener questionnaire by household size and sample type, 2009-10 BCS

	Single person	More than one person
	%	%
No-one present	93	62
Children under 16	1	11
Spouse/partner	*	24
Other adult	6	8
<i>Bases: All respondents</i>	<i>12,453</i>	<i>32,736</i>
Percentages add to more than 100% because respondents could give more than one answer		

The impact of the presence of others during the interview on the information given in the survey is not known as there is no way of knowing what the respondent might have said if they were alone. Table 4.3 shows the proportion of respondents who reported being a victim of crime by who was present during the screener survey. Respondents whose spouse or partner was present are less likely to report victimisation. However in cases where children under 16 are present or another adult is present respondents appear to be more likely to report having been a victim of crime.

It is likely however that other demographic factors may be influencing this such as age, gender, social behaviour etc)

Table 4.3 Reporting of victimisation by who else present during the screener questionnaire

	No-one present	Children under 16	Spouse/partner	Other adult	All households with more than 1 person
	%	%	%	%	%
Victim	24	31	20	31	25
Non Victim	76	69	80	69	75
<i>Base:</i>	<i>20,160</i>	<i>3,490</i>	<i>1,605</i>	<i>802</i>	<i>32,736</i>

Base: All with more than one person in the household

4.5.2 Presence of others during the self-completion and assistance given

For those who did the self-completion, the presence of others during this part of the interview was also recorded. Table 4.4 shows that almost three-quarters of respondents (73%) who did the self-completion did so when no-one else was present. Less than one in ten respondents (9%) who completed the self-completion did so when children were present in the room. As with the screener questionnaire, those who lived on their own were more likely than those who lived with other people to do the self-completion when no-one else was present (93% and 69% respectively).

Table 4.4 Whether anyone else was present or not during the self-completion by sample type, 2009-10 BCS

	Core sample
	%
No-one else	73
Spouse/partner/girlfriend/boyfriend	13
Child(ren) under 16	9
Other household member (adult)	6
Someone else	3
<i>Bases: All respondents who did the self-completion</i>	<i>26,899</i>
Percentages add up to more than 100% since more than one answer could be coded at this question	

Where anyone else was present in the room during the self-completion section, interviewers were briefed to try and 'arrange' the room whenever possible so that the respondent had a degree of privacy to do the self-completion. Thus, for example, interviewers might try to ensure that the respondent was sitting with the screen facing a wall or was in such a position that no-one else in the room could actually read the computer screen.

Where anyone else was present, the extent to which they were involved in answering questions was noted, as was whether the interviewer was involved in the self-completion sections. In cases where someone else was present during the self-completion, it was not common for others to become involved in answering the questions. In 89% of interviews where someone else was present, the respondent completed the questions entirely on their own. In 5% of interviews someone else actually looked at or read the self-completion with the respondent, while in another 6% of interviews the respondent discussed the self-completion with other people.

Respondents aged 45-59 (13%), Asian respondents (23%), and Black respondents (15%) were more likely than average to have had someone else look at or read the self-completion or to have discussed the self-completion with someone else.

Table 4.5 shows the amount of assistance that interviewers gave to respondents on the self-completion section. The vast majority of respondents who answered the questions (83%) used the laptop on their own without any help from the interviewer. About one in six respondents (17%) required some form of assistance with the self-completion. One in seven respondents (14%) asked the interviewer to enter their answers for them, while a further 3% of respondents entered their own answers but asked the interviewer for some help.

Respondents aged 45-59 (23%), Asian respondents (29%) and Black respondents (25%) were the most likely to have sought some help with the self-completion. This was primarily because these respondents were more likely to have asked the interviewer to complete the self-completion for them, rather than using the computer themselves.

Table 4.5 Amount of assistance given by interviewers with the self-completion questionnaire by sample type, 2009-10 BCS

	Core sample
	%
All done by respondent	83
Help given with one or two questions	2
Help given with more than one or two questions, but less than half	1
Help given with more than half, but not all	*
Help given with all/nearly all	1
Completed by interviewer	14
<i>Base: All respondents who did the self-completion</i>	<i>27,304</i>

4.6 Length of interview

4.6.1 Introduction

Timing stamps were placed throughout the questionnaire to allow timing of individual sections. Due to various technical issues associated with CAPI systems, it is not always possible to derive meaningful time stamps from every interview. For example, should an interviewer briefly go back into an interview at a later time to check or amend a response the time stamps can be set to show an apparently very short (2-3 minutes) interview. Similarly, if an interviewer has to temporarily stop or suspend an interview for an hour or so and fails to come out of the questionnaire in the intervening period (simply powering down the computer instead) the time stamps can show an interview of 4-5 hours.

To eliminate the effects of these outlying cases on the calculation of average timings, it was decided to only include interviews where the total length of interview was in the

range 15 minutes to 180 minutes. On the 2009-10 survey, around 99% of interviews had a valid time within these ranges and are included in the analysis below¹⁵.

4.6.2 Overall length of interview

The average (mean) core interview length during 2009-10 was 49 minutes. This is broadly the same length compared with recent years but has increased slightly since 2002 when the average length was 46 minutes. Table 4.6 shows the average interview length for the core sample since 2002.

Table 4.6 Average interview length over time

Survey Year	Average time (minutes)
2002-03	46
2003-04	46
2004-05	48
2005-06	48
2006-07	49
2007-08	48
2008-09	49
2009-10	49

The main influence on interview length was whether or not the respondent had been a victim of crime or not. The average interview length for non-victims was 46 minutes compared to 62 minutes for victims of crime.

The average length of interview by number of Victimization Modules is shown in Table 4.7 below. The length of interview was strongly related to the number of Victimization Modules completed by the respondent, with those completing 4 or more modules having an average interview length of 91 minutes.

¹⁵ Timings are based on the 2009-10 survey questionnaire which relates to all interviews conducted between 1st April 2009 and 31st March 2010.

Table 4.7 Average time of interview by number of Victimisation Modules, 2009-10 BCS

Number of Victimisation Modules	Average time (minutes)
Non victims	46
All victims	62
1	57
2	72
3	82
4 or more	92
All respondents	49

Most interviews took between 30 and 60 minutes, with 65% of all respondents completing the survey in this time. Just over one in eight (13%) completed the survey in less than 30 minutes, while 4% of respondents took 90 minutes or more.

Respondents aged 60 or over had a shorter average interview time compared with those aged under 60 (45 minutes and 52 minutes respectively), probably reflecting the fact that those aged 60 or over did not do the self-completion part of the interview and also the fact that older people are less likely to be victims of crime.

4.6.3 Average time for different sections of the interview.

The average times for each of the main modules of the questionnaire are shown in Table 4.8. It should be noted that this table shows the average times for each module across all respondents. Therefore, respondents who did not complete a particular module because of the sub-sampling are allocated a time of 0.

Table 4.8 Average time of each module of the questionnaire, 2009-10 BCS

Questionnaire module	Average time (minutes)
Household Box	3.4
Perceptions of crime module	7.1
Screeners questions	3.0
Victimisation Modules	3.2
Mobile phone and second home crime	0.5
Attitudes to the Criminal Justice System	8.0
Module A-D ¹	5.4
Plastic card fraud and Identity fraud	2.5
Anti Social Behaviour	1.0
Demographics	7.8
Drugs and Drinking	2.3
Inter-Personal Violence	2.0
End of interview administration	2.3
Average time	49

¹ This is the average time across all the four Modules A-D. Every respondent completes one of these modules and the aim is to ensure that they are as similar as possible in length.

4.6.3 Length of Victimisation Modules

As mentioned above the average length of the survey is affected primarily by the number of Victimisation Modules completed by a respondent, with the average time for non-victims being 46 minutes compared to an average of 62 minutes for victims of crime.

Although the average time taken to complete the Victimisation Modules was only 3.2 minutes across all respondents, this time is skewed by the fact that three-quarters of respondents were non-victims and so did not complete any Modules. Therefore, a more meaningful analysis is to look at the time taken to complete the survey by the number of Victimisation Modules completed.

Table 4.9 shows that long Victimisation Modules (1-3) averaged about 8 to 10 minutes per module, while short Victimisation Modules (4-6) averaged 4 to 5 minutes per module. The time taken to complete the first Victim Module was greater than for modules two or three, suggesting that respondents speed up as they go through each subsequent module. This pattern has been evident in all previous surveys.

Table 4.9 Average time of each individual Victimisation Module, 2009-10 BCS

Victim Module number	Average time (minutes)
Victim Module 1	10.0
Victim Module 2	8.7
Victim Module 3	7.9
Victim Module 4	4.6
Victim Module 5	3.9
Victim Module 6	4.1

4.6.4 Length of part-sample modules

Because the BCS survey is highly filtered each respondent only complete a certain number of modules. Table 4.10 below shows the average time taken for each of the part-sample modules based only on those respondents who were asked the module.

Table 4.10 Average time of different survey modules, 2009-10 BCS

Part-sample module	Average time (minutes)
Module A	5.1
Module B	8.0
Module C	5.0
Module D	3.6
Anti Social Behaviour (Modules B/D respondents)	2.0
Drugs and drinking self-completion	4.1
Inter-personal violence self-completion	3.6

The overall timings of the self-completion are masked by the fact that all those who are not eligible for the self-completion (i.e. those aged 60 years or over) and those who refuse the self-completion have an average time of zero. Considering only those respondents who actually did the self-completion sections, the average time of the Drugs and Drinking module was 4.1 minutes, while the average time of the Inter-Personal Violence module was 3.6 minutes.

Three quarters (75%) of respondents who completed the Drugs and Drinking module did it in under 5 minutes, while 1% of respondents took more than 10 minutes to complete it. For the Inter-Personal Violence module, 84% of those who completed it took less than 5 minutes, and 3% took more than 10 minutes.

4.7 Response rate and reasons for non-response: core sample

4.7.1 Overall core response rates

The full response rate analysis for the 2009-10 issued core sample is shown in Table 4.11.

One in ten issued addresses (10.6%) were identified as not being eligible residential addresses (known as deadwood). The most common type of deadwood was empty or vacant residential properties, which accounted for 6% of all issued addresses. The total proportion of addresses identified as deadwood has been gradually increasing over the last few years, but this trend seemed to accelerate in 2008-9. The increase is primarily driven by an increase in the recording of empty properties. This may be a result of a real increase in vacant property as a result of the recession with properties standing empty for longer before new residents move in. Alternatively it could be a result of interviewers becoming more skilled and experienced at identifying vacant properties over time. The importance of correctly identifying deadwood is emphasised in briefings and over time interviewers build up their own techniques for identifying vacant properties. It is also possible that there may be some mis-recording of non-contacts as deadwood by interviewers.

Interviewers made contact with either the selected respondent or a responsible adult at 97% of eligible addresses, meaning a non-contact rate of 3%. There were two types of non-contact. The most common (3% of eligible addresses) was where no contact was made with anyone at the address despite repeated calls over a lengthy fieldwork period. It is possible that some of these addresses were actually empty or vacant and so should have been coded as deadwood. However, the impact that this would have on the overall response rate is minimal. The remaining addresses classified as non-contact (0.5% of eligible addresses) were where contact had been made with someone at the address, but no contact was made with the person selected for interview.

At eligible addresses the most common reason for not getting an interview was due to a refusal, which accounted for 16% of all eligible addresses. The most common types of refusal were where the person selected for interview refused to take part in the survey (8%), and where no information about the household was given meaning that the person selection could not be carried out (4%). Proxy refusals (someone refusing on behalf of the selected respondent) and refusals directly to Head Office were less common.

A further 5% of eligible addresses were categorised as unproductive for other reasons including broken appointments, people who were ill or away during the period of the survey and people who had inadequate English to complete the survey.

Combining all the different types of unproductive addresses gave a final response rate of 76% for the 2009-10 survey. The response rate was similar to the previous year. In fact, response to the BCS has been broadly stable since 2001. Reasons for non-response were also broadly similar to previous surveys.

During the whole of 2009-10 a booklet of six first class stamps was sent with the advance letter as a 'thank you' to people for taking part in the survey¹⁶.

4.7.2 Performance against targets

Overall 45,189 interviews were achieved in 2009-10 against a target of 46,000 which is a shortfall of 811 interviews. The response rate achieved was 75.6%, slightly lower than the 76% target. Had the deadwood rate remained at the estimated level of 9% it is estimated that 45,998 would have been achieved with the response rate of 75.6%. However as the total proportion of deadwood addresses increased to 10.6% this resulted in a shortfall in the number of interviews achieved.

¹⁶ See Grant C. et. al. (2006) 2004/5 British Crime Survey (England and Wales) Technical Report (London: BMRB) for details of experiment carried out on BCS to test the impact of stamps on overall response rates.

Table 4.11 Core sample response rate and non-response outcomes, 2009-10 BCS

	N	% of issued addresses	% of eligible addresses
Total addresses issued	66,861	100.0	
Addresses not traced/inaccessible	394	0.6	
Not built/ does not exist	70	0.1	
Derelict/ demolished	279	0.4	
Empty/vacant	4,048	6.1	
Second home/not main residence	838	1.3	
Business/ industrial	1,099	1.6	
Institution/communal establishment	144	0.2	
Other deadwood	247	0.4	
Total ineligible addresses	7,119	10.6	
Total eligible addresses	59,742	89.4	100
No contact with anyone in household	1,758	2.6	2.9
No contact with selected respondent	282	0.4	0.5
Total non contact	2,040	3.1	3.4
Office refusal	1,082	1.6	1.8
Refused all information	2,562	3.8	4.3
Personal refusal	4,531	6.8	7.6
Proxy refusal	795	1.2	1.3
Contact made, no specific appointment	622	0.9	1.0
Total refusal	9,592	14.3	16.1
Broken appointment	721	1.1	1.2
Temporarily ill/incapacitated	196	0.3	0.3
Physically or mentally unable	762	1.1	1.3
Away/ in hospital	381	0.6	0.6
Inadequate English	373	0.6	0.6
Other unsuccessful	488	0.7	0.8
Total other unsuccessful	2,921	4.4	4.9
Total unproductive	14,553	21.8	24.4
Full interviews	45,140	67.5	75.6
Partial interviews	49	0.1	0.1
Total interviews	45,189	67.6	75.6

4.7.3 Core response rates by Government Office Region

Table 4.12 shows the different response rates and reasons for non-response achieved by Government Office Region in 2009-10. This shows that across most regions the response rate was broadly similar, ranging from 78% in North East to 75% in East of England. Only in London was response to the survey noticeably lower, with a final response rate of 67%. The lower response rate achieved in London was due to a slightly higher than average refusal rate (18%), but mainly due to a significantly higher non-contact rate (7%) compared with other regions. Lower response rates in London are a problem that is common to most major surveys, although the response achieved in 2009-10 was slightly higher compared with previous years (for example, 60% in 2005-06).

Table 4.12 Core sample response rates and non-response by Government Office Region, 2009-10 BCS

	Percentage of eligible addresses:			
	Non-contact	Refusal	Other unproductive	Achieved interviews
	%	%	%	%
North East	3.4	13.4	4.7	78.5
North West	3.1	16.0	4.4	76.5
Yorkshire & The Humber	3.9	16.1	5.3	74.7
East Midlands	3.1	15.3	5.3	76.2
West Midlands	3.1	15.3	4.4	77.3
East of England	2.6	17.4	4.8	75.2
London	7.4	18.5	6.8	67.4
South East	2.4	16.6	3.9	77.2
South West	2.1	15.4	4.4	78.0
Wales	3.8	15.2	5.1	75.8

4.7.4 Core response rate by Police Force Area

As outlined in section 2.5 the aim was to achieve around 1,000 interviews in each Police Force Area, with larger sample sizes in the most populous Areas. In order to achieve this sample size within each PFA the amount of sample issued was based on actual average deadwood rates and response rates over the period 2004-2007.

Table 4.13 below shows the actual number of interviews achieved in each PFA and the response rates. This shows that in a number of Areas the target number of achieved interviews exceeded 1,000, while in other areas the number of achieved interviews fell

slightly short. This is explained simply by the fact that the actual eligibility and response rates achieved in certain Areas in 2009-10 were slightly different (either higher or lower) from the figures used to estimate the amount of sample to issue.

Table 4.13 Core sample achieved interviews and response rates by Police Force Area, 2009-10 BCS

PFA	Number of interviews		Response rate %
	Target N	Achieved N	
Avon & Somerset	1,000	1,001	81.3
Bedfordshire	1,000	923	75.5
Cambridgeshire	1,000	1,007	75.5
Cheshire	1,000	990	78.3
Cleveland	1,000	963	77.2
Cumbria	1,000	968	78.3
Derbyshire	1,000	917	76.4
Devon & Cornwall	1,000	1,051	75.8
Dorset	1,000	926	75.2
Durham	1,000	1,040	79.3
Dyfed Powys	1,000	992	78.7
Essex	1,000	1,038	77.9
Gloucestershire	1,000	959	77.5
Greater Manchester	1,425	1,369	72.4
Gwent	1,000	922	71.5
Hampshire	1,000	1,048	79.5
Hertfordshire	1,000	956	72.1
Humberside	1,000	916	72.1
Kent	1,000	957	78.3
Lancashire	1,000	1,021	77.2
Leicestershire	1,000	949	74.0
Lincolnshire	1,000	1,032	80.6
Merseyside	1,000	928	78.7
Metropolitan	3,900	3,934	67.4
Norfolk	1,000	917	72.1
North Wales	1,000	980	81.8
North Yorkshire	1,000	999	74.3
Northamptonshire	1,000	1,022	76.3
Northumbria	1,000	1,041	78.9
Nottinghamshire	1,000	1,042	74.0
South Wales	1,000	969	71.8
South Yorkshire	1,000	967	76.7
Staffordshire	1,000	1,004	75.4
Suffolk	1,000	935	78.2
Surrey	1,000	919	73.6
Sussex	1,000	949	75.9
Thames Valley	1,125	1,112	78.4
Warwickshire	1,000	1,015	80.4
West Mercia	1,000	977	77.2
West Midlands	1,375	1,383	76.6
West Yorkshire	1,175	1,161	75.8
Wiltshire	1,000	987	80.7

4.7.5 Core response rates by type of area and type of property

Since large administrative areas such as Government Office Regions contain a variety of different types of area it is useful to examine response to the survey broken down by area type. Table 4.14 shows the response rates and reasons for non-response by different types of area. This shows that overall response rates tended to be lower in areas categorised as inner city compared with non inner city areas (69% and 76% respectively). This difference in response rate explains why the current BCS data includes a weight to correct for differential response rates between those areas defined as inner city and non-inner city (see [section 7.2.2](#)).

Similarly, the response rate in urban areas was lower compared with that achieved in rural areas (74% and 79% respectively). Response also varied significantly by ACORN Category, being highest in areas classified as 'Wealthy Achievers' (79%) and lowest in areas classified as 'Urban Prosperity' (69%). There was similar variation in response by Output Area Classification, ranging from 80% in 'Countryside' Areas to 67% in 'City living'¹⁷.

Looking at the differences in response rates by types of area shows how most of the response differential is due to variation in the non-contact rate, while the refusal rate tends to be fairly consistent. Thus, while the refusal rate varied between 14% and 18% in the different types of areas shown in Table 4.14, the non-contact rate varied from 2% to 8%.

¹⁷ For details of Output Area Classification see <http://areaclassification.org.uk/>

Table 4.14 Core sample response rates and non-response by types of area, 2009-10 BCS

	Percentage of eligible addresses:			
	Non-contact	Refusal	Other unproductive	Achieved interviews
	%	%	%	%
Inner city ¹	7.7	15.6	7.5	69.2
Non-inner city	3.0	16.1	4.6	76.3
Urban ²	3.8	16.4	5.3	74.4
Rural	2.2	14.9	3.5	79.4
ACORN Category				
Wealthy Achievers	2.0	15.9	3.1	79.0
Urban Prosperity	7.4	17.4	6.5	68.7
Comfortably Off	2.7	16.1	4.4	76.7
Moderate Means	3.4	15.8	5.6	75.1
Hard Pressed	4.5	15.7	6.6	73.2
Output Area Classification				
Blue Collar Communities	2.7	15.7	5.4	76.3
City Living	8.4	18.1	6.8	66.6
Countryside	1.8	14.5	3.4	80.3
Prospering Suburbs	2.0	16.7	3.3	77.9
Constrained by Circumstances	4.5	15.7	6.2	73.6
Typical Traits	3.3	16.4	4.5	75.8
Multicultural	7.0	16.0	8.5	68.4
¹ Inner city is based on the BCS definition that has been used for many years. See section 7.2 for more details.				
² This is based on the ONS definition of urban-rural areas, where urban is classed as 'urban – sparse' and 'urban –less sparse' and all other areas are classed as rural				

As mentioned in [section 4.4.2](#), part of the BCS assignment involved the interviewer collecting some details about the area and about the specific issued address. Since this information was collected for all residential addresses, whether or not an interview was obtained, it is possible to analyse response rates according to this data. Of most interest is how response varies first, by the type of property and second, by the type of area.

Table 4.15 shows how response rates on the 2009-10 survey varied according to the type of property, ranging from 81% among detached and semi-detached houses to 70% among flats.

The differential response rates achieved at different types of flats shows the impact on response rates of two particular aspects of flats, namely whether or not a property has a communal entrance and whether or not the communal entrance is lockable (e.g. controlled entry phone system). Not surprisingly, flats with communal entrances that had controlled entry systems were the most difficult type of property for interviewers to gain response. In 2009-10, the response rate at these types of property was 67% compared with 78% for flats with their own (non-communal) entrances. Flats with locked entrances had a higher than average level of non-contact (9%). This highlights the difficulty faced by interviewers in trying to gain an interview at an address where they are unable to make direct face-to-face contact with people, often having to communicate via intercom systems.

Table 4.15 Core sample response rates and non-response by types of property (recorded by interviewers), 2009-10 BCS

	Percentage of eligible addresses:			
	Non-contact %	Refusal %	Other unproductive %	Achieved interviews %
Detached/semi-detached house	1.8	13.7	3.9	80.6
Terraced house	3.2	13.6	5.4	77.7
Maisonette	4.0	13.2	5.5	77.4
<u>Flats with:</u>				
Own entrance	5.1	11.2	6.2	77.5
Non-lockable communal entrance	5.5	13.5	8.0	73.0
Lockable communal entrance	9.4	15.5	8.1	67.0
All types of flat	8.4	14.2	7.8	69.6

Taken together these figures go some way to explain the lower than average response rate in London, although there are clearly other factors involved as well. For the country as a whole, flats represented only 14% of the issued eligible sample, while flats with locked communal entrances represented 10% of the issued eligible sample. However, in London these types of properties represented 37% and 28% of the issued eligible sample respectively. Therefore, one important reason for the lower response rate in London, and inner city areas in general, is the composition of the housing stock and the greater difficulties faced by interviewers in making contact.

Apart from the actual type of property, interviewers were also asked to record their general observations about the area immediately surrounding each issued address with respect to a number of characteristics including how common rubbish or litter was, how common vandalism and graffiti was and how common run down houses were. These might be considered to be an indication of the degree of physical disorder within a particular area.

Although these observations are clearly open to a high degree of subjectivity, Table 4.16 shows that there was some association between interviewer observations and the final response rate. Response rates were highest in areas that had a low level of physical disorder and lowest in the areas that had the highest levels of physical disorder.

Table 4.16 Core sample response rate by evidence of physical disorder (recorded by interviewer), 2009-10 BCS

	Very common	Fairly common	Not very common	Not at all common
How common is...	%	%	%	%
Litter or rubbish lying around	74	75	77	80
Vandalism, graffiti or damage to property	78	75	77	79
Homes in poor condition or run down	76	74	78	79

4.8 Response to the self-completion questionnaire

The last part of the questionnaire involved a self-completion questionnaire which was asked of all respondents aged 16-59. In 2009-10 there were two self-completion modules on the survey:

- Use of illegal drugs and drinking
- Experience of domestic violence, sexual victimisation, and stalking.

Because of the sensitive nature of the questions these modules were asked using Computer Assisted Self Interviewing (CASI). Respondents were asked to complete the last part of the survey by entering the answers directly in to the laptop. Interviewers gave respondents a brief explanation of how to use the laptop, including taking them through some practice questions, before handing the laptop over. Interviewers were always present to help respondents if they needed any technical assistance. Once

respondents had completed the modules, they handed the laptop back to the interviewer.

Although respondents were encouraged to use the computer themselves, if they did not want to use it for some reason, interviewers were allowed to administer the modules provided that no-one else was present in the room. Where the self-completion part of the survey was administered by the interviewer the domestic violence, sexual victimisation and stalking modules were not completed, since these questions were considered too sensitive to be read out by the interviewer.

Table 4.17 shows that 93% of eligible respondents in the core sample answered the self-completion questionnaire, with 81% of them entering their answers directly in to the laptop themselves and 13% asking the interviewer to enter their answers for them.

Table 4.17 Response to the self-completion questionnaire, 2009-10

	Core sample
	%
Refused	7
Completed by interviewer	13
Accepted by respondent	81
Overall self-completion response	93
<i>Bases</i>	28,854

Table 4.18 shows how response to the self-completion questionnaire varied according to the demographic characteristics of respondents.

There was no difference between men and women in terms of response to the self-completion. Older respondents were slightly more likely than younger respondents to refuse to complete the self-completion questions (7% of 45-59 year olds compared with 5% of 16-24 year olds). More noticeable, however, was the fact that older respondents were more likely than younger ones to ask the interviewer to enter their answers for them (17% of 45-59 year olds compared with 7% of 16-24 year olds).

Some of the most noticeable differences were between respondents from different ethnic groups. Only 5% of White respondents refused to do the self-completion compared with 22% of Asian and 18% of Black respondents. Asian respondents were more likely than White respondents to ask the interviewer to enter their answers for them (18% of Asian respondents compared with 12% of White respondents).

There were also some differences by socio-economic classification, with respondents from routine and manual occupations being less likely than those from managerial and professional occupations to answer the self-completion (93% and 96% respectively). Respondents from routine and manual occupations were also more likely than those from managerial and professional occupations to ask the interviewer to enter their answers for them (18% and 8% respectively).

Table 4.18 Response to the self-completion questionnaire by socio-demographic characteristics of respondents (core sample), 2009-10 BCS

	Refused	Completed by interviewer	Accepted by respondent ¹	Overall self-completion response	Bases: N
	%	%	%	%	
Sex					
Male	7	13	80	93	13,116
Female	7	12	81	93	15,738
Age					
16-24	5	7	88	95	3,707
25-34	7	10	83	93	6,107
35-44	7	11	82	93	8,125
45-59	7	17	76	93	10,915
Ethnicity					
White	5	12	82	95	25,965
Mixed	7	11	82	93	290
Asian	22	18	61	78	1,286
Black	18	12	69	82	753
Other ethnic group	21	14	65	79	517
NS-SEC					
Managerial & professional occupations	4	8	88	96	9,663
Intermediate occupations	6	13	81	94	5,344
Routine & manual occupations	7	18	75	93	9,547
Unclassified	12	11	76	88	2,750
Total	7	13	81	93	27,304

¹ Respondent used the laptop on their own

Table 4.19 shows the reasons given by respondents either for refusing the self-completion questionnaire or for asking the interviewer to enter their answers for them. This shows that a dislike of computers was the most common reason why respondents asked the interviewer to enter their answers for them (mentioned by 38%), while running out of time was the most common reason given for respondents refusing to do it (mentioned by 55%). Language problems were a reason given by 12% of respondents who refused the self-completion and 6% of those who asked the interviewer to do it for them. Only 5% of respondents refused to do the self-completion questionnaire because of worries about confidentiality.

Table 4.19 Reasons for refusing self-completion questionnaire or for completion by interviewer (core sample), 2009-10 BCS

	Refused	Completed by interviewer	Total
	%	%	%
Don't like computers	12	38	29
Ran out of time	55	30	39
Language problems	12	6	8
Couldn't be bothered	7	9	8
Children in room	9	8	8
Disability	4	5	4
Eyesight problems	2	4	3
Could not read/write	2	3	3
Confidentiality worries	5	1	2
Other people in room	3	1	2
Objected to study	2	1	1
Other reasons	7	8	8
<i>Bases:</i>	<i>1,922</i>	<i>3,634</i>	<i>5,556</i>
Percentages add up to more than 100% since more than one answer could be coded at this question			

Table 4.20 shows the reasons given by people who refused the self-completion or who had the interviewer enter their answers for them broken down by age and ethnic group. This shows that older respondents were more likely than younger respondents to cite a dislike of computers as a reason for not doing the self-completion or for asking the interviewer to enter their answers for them (mentioned by 41% of 45-59 year olds compared with 11% of 16-29 year olds). Respondents aged 45-59 were also slightly most likely to mention eyesight problems (5%) and having a disability (6%). Respondents aged 16-29 were more likely than older respondents to cite the fact that children were present in the room as a reason for refusal or asking the interview to enter their answers for them (18%) or language problems (13%).

Non-white respondents were more likely than white respondents to mention language problems as a reason for refusing the self-completion or asking the interviewer to enter their answers for them. This was given as a reason by 38% of Asian respondents and 11% of Black respondents.

Table 4.20 Reasons for refusing self-completion questionnaire or for completion by interviewer by age and ethnic group (core sample), 2009-10 BCS

	Age			Ethnic group				
	16-29	30-44	45-59	White	Mixed	Asian	Black	Other
	%	%	%	%	%	%	%	%
Ran out of time	43	44	34	39	44	36	45	33
Don't like computers	11	20	41	31	14	19	19	14
Children in room	18	12	2	8	12	8	11	6
Couldn't be bothered	8	8	8	8	10	7	15	4
Language problems	13	11	4	4	2	38	11	43
Disability	3	4	6	5	8	1	4	2
Eyesight problems	2	1	5	3	0	1	3	2
Could not read/write	3	3	2	2	4	5	3	3
Confidentiality worries	2	2	2	2	2	3	6	3
Other people in room	4	2	1	2	2	2	0	3
Objected to study	1	2	1	2	0	1	1	1
Other reasons	7	7	8	8	8	5	8	6
<i>Bases:</i>	898	2,011	2,647	4,581	50	507	231	181

Percentages add up to more than 100% since more than one answer could be coded at this question

4.9 Full and Partial Interviews

For an interview to be regarded as valid, respondents had to answer to the end of the screener questions. Any interview which was abandoned before the end of the screener questions was not regarded as useable and was not put on the data file.

An interview was counted as a full interview for the core sample if the respondent completed to the end of the demographics module. If the interview was stopped before the end of the demographics module it was coded as a partial interview. Full and partial interviews were recorded separately in the field figures. In 2009-10, 99.9% of interviews achieved on the core sample were full interviews and only 0.1% (n=49) were partial interviews.

5. Data Processing

5.1 Offence coding

The BCS Offence Coding System was developed for the 1982 BCS to match as closely as possible the way incidents were classified by the police. This involves collecting detailed information about incidents reported by respondents in the Victimization Modules. Once the data is returned to the office, all Victimization Modules are reviewed by specially trained coders in order to determine whether what has been reported represents a crime or not and, if so, what offence code should be assigned to the crime.

Apart from some minor changes, the code frame and the instructions to coders have remained stable since 1982. The operational procedures used for assigning codes on the 2009-10 survey have been in place since 2001.

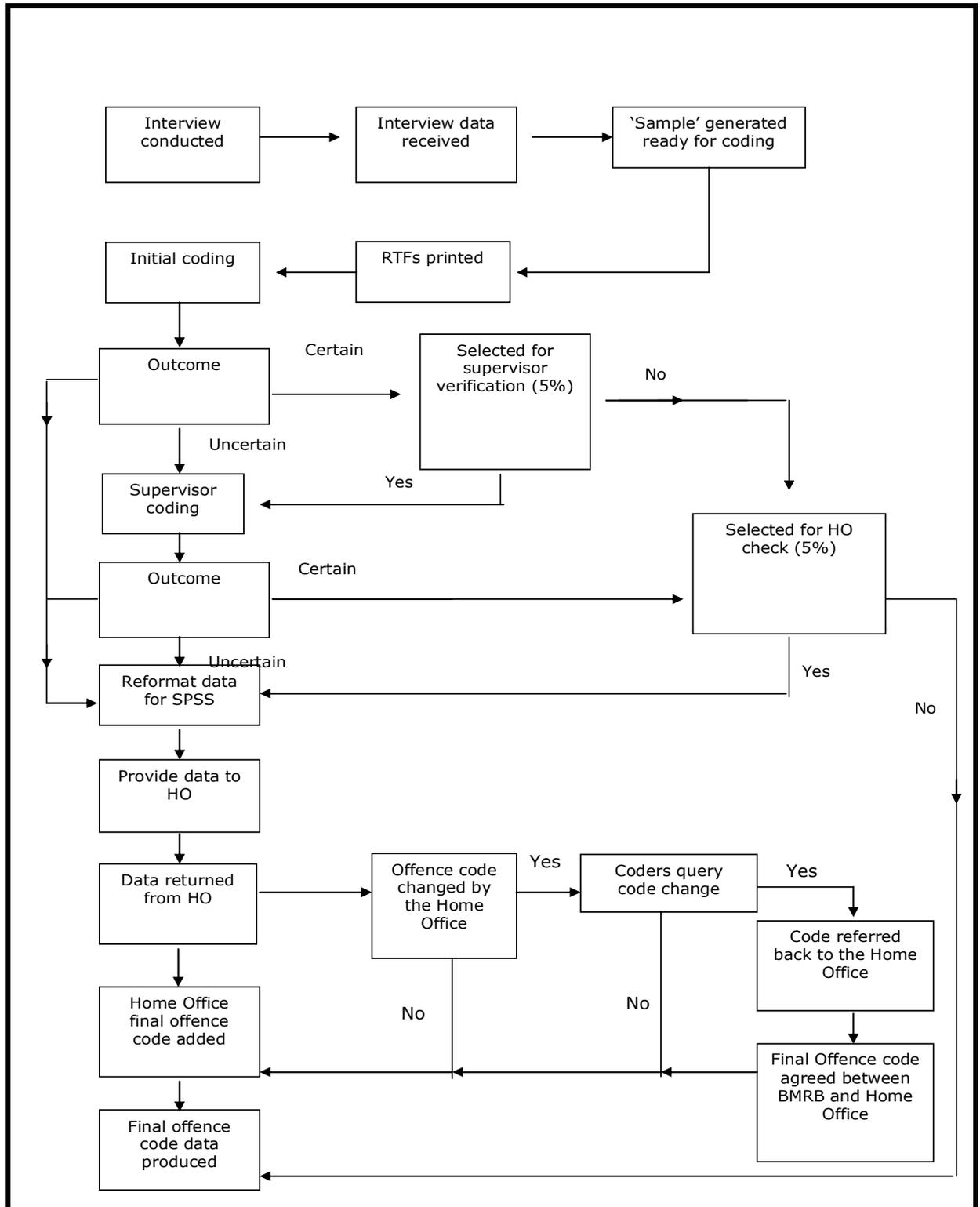
During 2009-10, the Offence Coding System consisted of the following steps:

- For each Victimization Module a paper-based summary was produced. This represented the key information from the CAPI questionnaire, including the verbatim description, which coders needed to enable them to assign an offence code.
- In addition to these paper-based summaries the coders used a specially developed computer assisted questionnaire to help them arrive at a final offence code for each Victimization Module. The questionnaire is written in a way that allows coders to arrive at a particular offence code by answering a series of questions using the information they have available on the paper-based summary. Additionally coders have a full reference manual (see Appendix G of Volume 2). It should be stressed however, that the computer assisted questionnaire is simply a tool to help coders arrive at a final offence classification. The final decision ultimately relies on the skill and judgement of the coder.
- A supervisor checked any codes that the original coder was uncertain about. Additionally, 5% of codes where the coder was certain of the outcome were also checked as a further quality check. These are systematically selected from all cases that have been coded (i.e. every *n*th case) in a particular period.

- Researchers at the Home Office checked:
 - Any codes that TNS-BMRB were uncertain about
 - Certain types of incident that were automatically referred (e.g. arson)
 - A proportion of certain codes as part of a quality control check

The result of this process was that every Victimisation Module had a final offence code assigned to it. A flow chart of the Offence Coding System is shown in Figure 5.1 and the offence coding system is explained in more detail below.

Figure 5.1 British Crime Survey Offence Coding Flowchart



5.1.1 The automatically generated offence code

In 1996 a programme was introduced that automatically generated an offence code based on the answers to a number of pre-coded variables in the Victimisation Module. The programme that was used for the 2009-10 survey was the same as that used on the survey since 2001.

An automatic code cannot be generated in all cases, and in 2008-09 no automatically generated code was produced for just over a quarter (27%) of Victimisation Modules. Reasons for this included missing codes or due to some inconsistency between the different variables used. Irrespective of the suggested automatic code, the coder has the responsibility of producing an offence code, and coders are instructed to see the generated code as only a starting point.

On the 2009-10 survey for Victimisation Modules where a code was automatically generated, it was the same as the final offence code in 75% of cases.

5.1.2 The coding task

Coders are provided with a paper-based print out of the key variables from each Victimisation Module and this information forms the basis of the coding. This document also provides coders with the offence code that had been generated by the automatic generation programme. An example of this paper form can be found in Appendix G in Volume 2.

Coders used a specially designed computer assisted questionnaire to carry out the coding. The questionnaire asked the coders certain questions about the nature of the offence. The questionnaire takes account of the major rules that apply to offence coding (such as the priority of codes), and by answering the questions on the basis of the information provided in the Victimisation Module, the coders reach an offence code.

All coders were personally briefed about the offence coding. The coders were also provided with a coding manual. This manual is similar to the one used in previous years of the BCS and contains all the rules that govern offence coding. The manual also provides flow-charts that show how the coding questionnaire works, so that coders can see how they reached a particular offence code on the basis of the

answers that they input. A copy of this manual is provided in Appendix G in Volume 2.

When the coder reaches an offence code, they can say whether they are certain or uncertain that this is the right code. Any Victimization Module which the coder is uncertain about is automatically referred to their supervisor for checking. In addition, the supervisor checks 5% of codes which coders were certain about.

5.1.3 Home Office coding

All cases where the coders are uncertain about the correct code to assign are automatically referred to the Home Office.

In addition to this, 5% of all codes which TNS-BMRB were certain about were selected to be sent to the Home Office for quality control checking. These were selected in a systematic fashion by selecting every *n*th case in each two-week time period.

A list of Victimization Modules to be checked by researchers at the Home Office was sent every two weeks. This consisted of an Excel spreadsheet that contained the unique serial number of each Victim Module, the code that the coder (and supervisor if applicable) had given the incident, how certain the coder (and supervisor) was about the coding, and any notes that the coder added about why they were uncertain. An electronic version of the paper-based document providing the key variables from the Victimization Module was also provided.

Researchers at the Home Office coded each of the Victimization Modules sent to them (using the paper-based document) and returned the spreadsheet with their code and any comments added. These codes were then manually added into the coding file (so that the coders could see the changes that had been made).

Particular attention was paid to cases where the Home Office changed a code that TNS-BMRB coders had marked as “certain”. If the TNS-BMRB coders disagreed with such a coding decision, this was fed back to both TNS-BMRB researchers and Home Office researchers for further consideration and discussion.

In total 1,558 cases were sent to the Home Office for checking as part of the 2009-10 survey, which represented about 11% of all Victimization Modules.

Of the Victimisation Modules sent to the Home Office:

- 22 were code 01s which were automatically referred to Home Office. This covers cases of aggravated burglary, duplicate cases and cases where the Victimisation Module was invalid;
- 173 were code 02s (cases where the TNS-BMRB coder was not certain about the code) which were also automatically referred to the Home Office for checking.
- 661 were part of the quality control check.
- 702 were related Victimisation Modules. To ensure that those checking offence codes had complete information all the Victimisation Modules belonging to an individual respondent were sent to the Home Office, rather than just the single Module under consideration.

Of the 1,558 Victimisation Modules sent to the Home Office 142 cases had their code changed by the Home Office, representing 9% of all cases sent. This level of change was fairly static across the survey year suggesting a degree of stability in the offence coding process.

The codes changed by the Home Office according to the categories outlined above were as follows:

- in 6 cases offences were coded 01 for referral to the Home Office. As code 01 is not a valid this code was changed in all cases;
- in 9 cases where the module was judged to be invalid by TNS-BMRB coders, 3 (33%) were changed by the Home Office;
- in 7 cases referred as duplicates no codes were changed
- in 175 cases where TNS-BMRB coders were uncertain, 57 (33%) were changed by the Home Office;
- in 661 cases sent for quality control 32 (5%) were changed by the Home Office; and
- in 702 related cases, 44 (6%) were changed by the Home Office

In all cases where the Home Office changed a code that TNS-BMRB coders or supervisors had been certain about, this was double checked and verified by TNS-BMRB upon return of the coding from the Home Office. Where TNS-BMRB did not

agree with the Home Office decision cases were referred back to the Home Office for re-checking. Of the 142 cases changed by the Home Office, 34 were referred back for re-checking. In 18 cases the original TNS-BMRB code was deemed to be correct and was re-instated as the final code and in 10 cases the Home Office code was deemed to be correct. For the remaining 6 cases a different code was decided upon after further discussion. After all queries had been resolved 124 cases were changed by the Home Office, representing 8% of all cases sent.

5.1.4 Final Offence Code

The SPSS data set delivered to the Home Office includes all the offence codes that have been given to each Victimisation Module at every stage of the coding process. This allows a complete history of each case to be maintained at all times. The final offence code is derived using a priority ordering system, whereby the Home Office code takes priority over the supervisor code, which takes priority over the original coder code. The variables supplied to the Home Office are:

OFFSUG	Suggested offence code (generated by computer)
VOFFENCE	Code assigned by the original coder
SOFFENCE	Code assigned by the supervisor
FINLOFFC	Code assigned by the Home Office research team
OFFENCE	Final offence code

5.1.5 Checks on final offence code

During the creation of the SPSS data sets some further consistency checks are run on the final offence codes, checking these against key pre-coded variables in the Victimisation Module. The purpose of this is to highlight cases where some of the pre-coded data seems potentially anomalous with the final offence code. Such anomalies can arise because sometimes the information reported by the respondent is not consistent. In particular, there may be inconsistencies between the verbatim description of the incident and subsequent pre-coded questions. While interviewers are carefully briefed to try and be aware of such inconsistencies arising during the interview it is inevitable that some will be missed. Furthermore, consistency checks within the actual questionnaire script to try and pick up anomalies are not possible when a verbatim description is involved.

The consistency checks carried out are as follows:

- Assaults where no force or violence was recorded as having been used
- Burglary where entry to the property was recorded to be authorised
- Car thefts where no car was recorded as being stolen, or where the police were not informed
- Sexual assaults where there was no sexual element to the assault recorded
- Snatch thefts where the item stolen was not recorded as being held or carried
- Other thefts where the item stolen was recorded as being held or carried
- Wounding where no injury was recorded as being sustained
- In scope offences where the offender was perceived by victim to be mentally ill
- Thefts where nothing has been recorded as having been stolen
- Vandalism where no damage has been recorded
- Threats where no threat has been recorded

All cases that fail these checks are examined individually by a researcher and, if necessary, referred to the Home Office. Where clear anomalies in the data do exist it is up to the judgment of the researchers to decide which bits of information should be prioritised in arriving at the final agreed offence code. In such cases, greater credence tends to be given to a good verbatim description of the incident over the answers to specific pre-coded questions where for example anomalies may be a result of interviewer mis-keying.

Experience of running these checks shows that most flagged cases do have the correct offence codes, but a few may be amended each quarter as a result of this additional check.

5.2 Other coding

In addition to the Offence coding, coders also looked at all questions where an “other –specify” had been given as an answer. The aim of this exercise, commonly known as back coding, was to see whether the answer given could actually be coded into one of the original pre-coded response options. Coding was done in Ascribe, a Windows based coding package.

Coders were provided with the code frames used in the questionnaire as a starting point. Since most of the questions have been used in previous years of the survey, the code frames were already well developed and there was little need to add new codes to the frames. However, if the coding supervisor felt an extra code was needed, this was flagged up to researchers who approved any changes before they were implemented.

5.3 Coding of occupation and socio-economic classification

Occupation details were collected for all respondents, either relating to their current job or to their last job if the respondent was not currently employed but had worked at some time in the past. Occupational details of the Household Reference Person were also collected, if this was not the same person as the respondent.

Occupations were coded using the Standard Occupational Classification 2000 (SOC2000). All occupational coding was done centrally by specialist coders once the data were returned by interviewers. Coding was done using CASCOT, a package widely used to code occupation, with coders using the manuals for reference.

As well as occupation codes, National Statistics Socio-Economic Classification (NS-SEC) was added to the file for all respondents and Household Reference Persons. NS-SEC categories were derived automatically using an algorithm which was developed from the documentation provided by ONS. Both the NS-SEC operational categories and the NS-SEC analytical categories were derived.

Details of the NS-SEC categories can be found in Appendix H of Volume 2.

6. Data Output

6.1 Introduction

The main outputs provided to the Home Office are SPSS data files that are delivered on a quarterly basis. Separate data files are provided for the core sample and the under 16s sample. For each type of sample, two data files are provided: the Non Victim File and the Victim File.

The **Non Victim File (NVF)** is produced at the level of the individual respondent and contains all questionnaire data and associated variables, except for information that is collected in the Victimisation Modules. Data for both victims and non-victims are included on the Non Victim File.

The **Victim File (VF)** is produced at the level of the individual incident or crime and contains all the data collected in the Victimisation Modules. Thus, an individual respondent who reported three crimes and completed three Victimisation Modules would have three separate records in the Victim File. All generated Victimisation Modules were included on the file, including cases where the module either had been suspended or where the reference period was out of scope. Although such records contain no information and are not used for analysis, it is useful to keep these on the file to monitor the number of modules that fall into these categories.

6.2 Delivery of data output

During 2009-10 data files were supplied to the Home Office on a quarterly basis. Data was supplied on a 12 month rolling basis, meaning that each new data delivery was updated by adding the newest quarter of data and deleting the oldest quarter of data. In addition to the achieved sample, a data file of the entire 2009-10 issued sample was supplied to the Home Office. This contained information on every issued address such as the final outcome, the screening outcomes, the observational data collected by interviewers, sample variables and geo-demographic variables.

Data was delivered to the Home Office approximately five weeks after the end of each quarterly fieldwork period. Each quarterly data delivery included interviews that were **achieved** in a specific 12 month period, rather than those that were **issued** in a

specific time period. Thus, the four sets of quarterly data files delivered in 2009-10 covered all the relevant interviews achieved in the following periods:

- July 2008 – June 2009
- October 2008 – September 2009
- January 2009 – December 2009
- April 2009 – March 2010¹⁸

6.3 Content of SPSS data file

The SPSS data files delivered to the Home Office contain various types of variables. The main types of variables contained on the files are:

- **Questionnaire variables** (NVF and VF).
- **Geo-demographic variables** (NVF only). All interviews had a set of pre-specified geo-demographic variables attached to them (see Appendix H in Volume 2 for complete listing).
- **Observational variables** (NVF only). All interviews had the observational data collected by interviewers on the Address Contact Sheets attached to them (see Appendix H in Volume 2 for complete listing). Due to the way in which the Observational data was processed it was difficult to do this on a quarterly basis. Consequently it was agreed that Observational variables only be supplied on the main data set on an annual basis, as well as being supplied on the issued sample file mentioned in section 6.2.
- **Coding variables** (NVF and VF). On the Non Victim File, SOC2000 codes are included for both the respondent and the Household Reference Person. Additionally, NS-SEC for both the respondent and the Household Reference Person are included. On the Victim File, a full set of offence codes are attached as outlined in section 5.1.5.
- **Derived variables** (NVF and VF). Many derived variables were also added to the file. These consisted primarily of 2 types:
- **Flag variables** that identify, for example, the type of sample, the part-sample module split and sub-split, the date of interview, the month of issue, whether a

¹⁸ The April 2008 – March 2009 data file is the data on which the main 2009-10 annual crime figures are based and is the basis of the file deposited by the Home Office at the UK Data Archive.

partial or full interview, whether a victim or non-victim, etc. On the Victim File, flag variables include whether the record was a Long or Short Victimization Module, whether it was a Series or a Single incident, and whether it was inside or outside the reference period.

- **Classificatory variables** derived from the data. These included standard classifications such as ONS harmonised variables, banded age groups, ethnic groups, income groups, etc.
- **Weighting variables** (NVF only).

6.4 Conventions used on SPSS Data Files

In creating the 2009-10 data files great attention was paid to ensuring as much consistency as possible was maintained with previous years of the survey.

6.4.1 Case identifier

The case identifier was required to be similar to that used on previous years of the survey but also had to be designed to meet the requirements of a continuous survey. On the Non-Victim File, where each individual case or record represents an individual respondent, the unique case identifier (ROWLABEL) is an 8-digit number constructed as follows:

	Column position	Values
Year of issue	1	1-9
Area point number	2-5	1000-9999
Address number	6-7	01-32
Screen number ¹⁹	8	0-9

On the Victim File, where each individual case or record represents a Victimization Module or incident, the unique case identifier (MATCH) is a 9-digit number, which is identical to ROWLABEL with the addition of the Victimization Module number:

¹⁹ Screen numbers are used to identify the type of sample. '0' indicates a core sample case and '9' indicates a young adult boost case.

	Column position	Values
Year of issue	1	1-9
Area point number	2-5	1000-9999
Address number	6-7	01-32
Screen number	8	0-9
Victimisation Module number	9	1-6

6.4.2 Naming conventions

Variable names were kept the same as on the previous surveys wherever possible. Consistency is particularly important on a continuous survey where data from one survey year is combined with data from a previous survey year as described in section 6.2. However, this means it is also important to systematically document changes to questions over time to avoid confusion amongst users. For example, small changes to a question from one year to the next (such as adding an extra code to the code frame) can create the possibility of wrongly merging data that appears similar but, in fact, is not. To avoid such situations, the variable names on the 2009-10 data file were changed to reflect any variables where such changes had been introduced between 2008-09 and 2009-10. A list of variables which changed slightly between 2008-09 and 2009-10 is shown in Table 6.1.

Table 6.1 Changes in variables between 2008-09 and 2009-10 survey

Non Victim File			
Module	2008-9 variable	2009-10 variable	Reason for change
PERC	NOWALK2A-O	NOWALK3A-R	Change of code frame
CJS	COMMAT4A-Q	COMMAT5A-U	Change of code frame
FUC	YSECYR3A-W	YSECYR4A-W	No additional codes, but wording of existing codes changed
FUC	YNOALARA-L	YNOALA2A-M	Change of code frame
FRAUD	QLOSS	QLOSS2	Change of code frame
FRAUD	ID1A-J	ID1AA-L	Change of code frame
FRAUD	IDKNOWA-M	IDKNOW2A-N	Change of code frame
FRAUD	ID10	ID10A	Change of code frame
DEMOG	ETHNIC	RESETH RESETH01 RESETH02 RESETH03 RESETH04	Question format and position on survey changed
DEMOG	TOTHHIN1 INDINC	TYPEINCA-Q PERSINC OTHINC TOTHHIN2	Way in which income collected changed.
Victim File			
VF	HATEMOTA-G	HATEMT2A-H	Change of code frame
VF	HATEPOSA-G	HATEPS2A-H	Change of code frame
VF	WHYHAP2A-R	WHYHAP3A-S	Change of code frame
VF	WHATDO	WHATDO2	Change of code frame
VF	RACEOFF2	RACEOFF3	Change of code frame
VF	RACEOF3A-G	RACEOF3A-H	Change of code frame
VF	WHAST4A-SS	WHAST5A-UU	Change of code frame
VF	YMOBNRPA-H	YMBNRP2A-I	Change of code frame
VF	WHTRS4A-SS	WHTRS5A-UU	Change of code frame
VF	HTRYCAR	HTRYCA2	Change of code frame
VF	ALCCHE	ALCCHEB	Question wording changed
VF	TRYPRE2A-R	TRYPRE2A-T	Change of code frame
VF	HOWCTOL3	HOWCTOL4	Change of code frame

6.4.3 Labelling variables

The changing nature of the 12-month reference period over the course of the year creates a difficulty in labelling certain variables. In the Quancept script, dates were automatically calculated based on the date of interview and appropriate text substitution was used to ensure that the question always referred to the correct period. In the SPSS data files, which contain data from interviews achieved over the whole year, it is difficult to attach meaningful labels to certain variables since the label is different each month depending upon the month of interview. This issue affects the following variables (all on the Victim File):

- DATESERA-DATESERH
- NQUART1-NQUART5
- QTRRECIN
- QTRINCID

Details of how the code frames for these specific questions relate to the month of interview can be found in Appendix F of Volume 2.

6.4.4 Don't Know and Refused values

The convention for Don't Know and Refusal codes used in the most recent surveys was maintained on the 2008-09 data. This meant that on the SPSS file the code for Don't Know was '9' for code frames up to 7, '99' for code frames up to 97, and so on. The code for Refused was 8, 98, and so on. Since these are standard codes used throughout the SPSS files, Don't Know and Refused codes are not labelled.

6.4.5 Multiple response variables

Prior to the 2001 survey, multiple response variables were created as a set of variables equal to the maximum number of answers that could be given. The first variable held the first answer given by the respondent; the second variable held the second answer given, and so on.

After discussions with the Home Office it was agreed from 2001 onwards to present multiple response variables differently from previous years. Instead, multiple response variables were set up as a set of variables equal to the total number of answers possible (including Don't Know and Refused). Each variable was then given a value of '0' or '1' depending on whether the respondent gave that particular answer or not. To denote this change all multiple response variables in 2001 were all named

with a letter suffix, rather than the number suffix that was used in previous years of the survey.

An example of a multiple response variable where there are seven possible answer categories, and so seven separate variables, is shown below:

**AGEOFFA-
AGEOFFG**

[ASK IF NumOff IN (2..4)]

How old were the people who did it? Would you say they were...READ OUT
CODE ALL THAT APPLY

- | | | |
|----|-------------------------------|-----------|
| 1. | children under school age | (AGEOFFA) |
| 2. | children of school age | (AGEOFFB) |
| 3. | people aged between 16 and 23 | (AGEOFFC) |
| 4. | people aged between 25 and 39 | (AGEOFFD) |
| 5. | or people aged over 40? | (AGEOFFE) |
| | Don't Know | (AGEOFFF) |
| | Refused | (AGEOFFG) |

7. Weighting

7.1 Overview of weighting

The following weights have been calculated for the 2009-10 BCS data:

- A household weight for the core sample
- An individual adult weight for the core sample

In addition to these weights, the Home Office apply additional calibration weights once they receive the data so that the data reflect the population profile by age and sex within Government Office Region (see [section 7.4](#)).

There are three main reasons for computing weights on the BCS:

- To compensate for unequal selection probabilities. In the BCS, different units of analysis (households, individuals, instances of victimisation) have different probabilities of inclusion in the sample due to factors such as over sampling of small Police Force Areas, the selection of one dwelling unit at multi-household addresses, the selection of one adult in each household, and the inclusion of a single Victimisation Module to represent a series of similar incidents.
- To compensate for differential response. Differential response rates can arise both between different geographic units (e.g. differences in response between inner city and non-inner city areas) and between different age and gender sub-groups.
- To ensure that quarters are equally weighted for analyses that combine data from more than one quarter.

As outlined above a variety of different weights were computed to meet the different analysis requirements. The 2009-10 weighting schedule was broadly similar to the weighting schedule applied on previous surveys.

All weights include a component to compensate for unequal selection probabilities, while weighting components to compensate for differential response and to equally weight quarters are included in some weights but not in others.

7.2 Component weights

The weights constructed for the 2009-10 BCS sample were based on a number of key component weights. In constructing all the different weights the following conventions have been used for the components that made up the final weights:

- w_1 : weight to compensate for unequal address selection probabilities in each PFA;
- w_2 : inner city versus non inner-city non-response weight;
- w_3 : dwelling unit weight;
- w_4 : individual selection weight;
- numinc : series of incidents weight

7.2.1 Police Force Area weight (w_1)

Under the survey design introduced in 2008-09 the address sampling probability is a function of the Police Force Area, the cluster stratum and, in a few cases, the number of addresses sampled within the PSU. These can be explained as follows:

1. **Police Force Area**: As already described in Chapter 2, addresses were disproportionately sampled in Police Force Areas to ensure a minimum of 1,000 achieved interviews in each Area regardless of the population size. Consequently the basic sampling fraction applied within each PFA varies significantly between different Areas;
2. **Cluster stratum**: As already explained in Chapter 2 all addresses were allocated to one of three cluster strata. While the intention was to allocate proportionately, the requirement to sample whole number PSUs within cluster strata B and C lead to a tiny level of between-strata variation in address sampling probabilities. This could have been corrected by altering the number of addresses selected within each sampled PSU, but this was not done. Instead a standard number of addresses (32) were issued in each PSU sampled from strata B and C; and
3. **The number of addresses within the PSU**: A small number of very large PSUs had a computed sampling probability greater than 1. This is because the size of the PSU (as measured by the PAF address count) was larger than the selection interval, meaning they had a 100% chance of selection. In this situation the PSU sampling probability was capped at 1 but the number of addresses sampled within these PSUs was not increased to compensate for

this. This introduced another slight variation in address sampling probabilities. Only a handful of PSUs were affected by this.

While the above represents a full explanation of the address sampling probability it is only the Police Force Area which actually introduces any significant variation in probabilities. Factors 2 and 3 above only introduce extremely minor variations in probabilities within each PFA. Consequently, it is probably easiest to think of w_1 as the Police Force Area weight, which compensates for different selection probabilities between Areas.

7.2.2 Inner city weight (w_2)

In some previous rounds of the BCS, inner city areas were over sampled meaning that an inner city weight was applied. Historically this weight compensated not only for the difference in selection probabilities but also for the differential response rates between inner city and non-inner city areas.

To be consistent with previous survey years the practice of applying a weight to correct for differential response rates between inner city and non-inner city areas has continued. In essence, the inner city weight is simply the reciprocal of the achieved response rate in inner city and non-inner city areas (after weighting by w_1).

The definition of inner city or non-inner city has been kept consistent since it was first used on the BCS and is based on 1981 census data. Details of how the inner city weight is constructed can be found in previous BCS Technical Reports.

7.2.3 Dwelling unit weight (w_3)

At addresses which had more than one dwelling unit, the interviewer made a random selection of one dwelling unit. The dwelling unit weight is therefore simply the number of dwelling units identified at the address. In over 99% of cases, the dwelling unit weight was 1.

7.2.4 Individual weight (w_4)

At dwelling units that had more than one eligible adult, the interviewer made a random selection of one adult. Thus, the probability of any one individual being

selected was inversely proportional to the number of adults in the household. The individual weight is therefore simply the number of adults in the household.

7.2.5 Series weight (numinc)

This weight is applied when estimating victimisation rates. For single incidents NUMINC is always 1. For series incidents, where only details are collected about the most recent incident in the series, the weight equals the number of incidents in the series that fall within the reference period, subject to a maximum limit of 5²⁰.

In estimating victimisation rates, the household or individual weights are multiplied by the NUMINC weight, according to which offence classification code has been assigned to the incident(s).

7.3 Core sample weights

The main units of analysis used on the BCS are households, individuals, and incidents of victimisation. Different weights are used depending upon the unit of analysis. In particular, some crimes are considered household crimes (e.g. burglary, vandalism to household property, theft of and from a car) and therefore the main unit of analysis is the household, while others are personal crimes (assault, robbery, sexual offences) and the main unit of analysis is the individual.

For the core sample two weights were constructed to take account of this difference, namely the **core household weight** and the **core individual weight**. These were calculated as follows:

$$\mathbf{wtm2hhu} = w_1 * w_2 * w_3$$

$$\mathbf{wtm2inu} = w_1 * w_2 * w_3 * w_4$$

Once the unscaled weights had been calculated the frequencies were examined and extreme values were capped where necessary. Although capping of extreme weights may introduce a small amount of bias this is more than compensated for by

²⁰ Although the number of incidents is capped at 5 for weighting purposes, the actual number of reported incidents in each series (uncapped) is also supplied on the data file

the improvement in precision that results. The capped weights were called **wtm2hhf** and **wtm2inf** respectively.

Finally, the weights were scaled to a notional sample size of 11,500 interviews per quarter. Although an approximately equal number of addresses were issued each quarter during 2009-10, the number of interviews actually achieved per quarter varied to some extent. Thus, for analyses based upon a 12 month period, the weights were constructed to adjust for differences in sample size by equalising the quarterly achieved sample sizes.

The final scaled weights were called **wtm2hhs** and **wtm2ins** respectively.

7.4 Calibration Weights

From 2001 onward the Home Office have calculated and applied additional calibration weights to counter the effect of differential response rates between age, gender and regional sub-groups. Results for BCS surveys from 1996 onwards have all been re-weighted using this technique²¹.

Calibration weighting is designed to make adjustments for known differentials in response rates between different age by gender subgroups and households with different age and gender composition. For example, a 24 year old male living alone may be less likely to respond to the survey than one living with a partner and a child. The procedure therefore gives different weights to different household types based on their age and sex composition in such a way that the weighted distribution of individuals in the responding households matches the known distribution in the population as a whole.

The effects of applying these weights are generally low for household crime, but are more important for estimates of personal crime, where young respondents generally have much higher crime victimisation rates than average, but also lower response rates to the survey. However, crime trends since the 1996 survey have not been altered to any great extent by the application of calibration weights.

²¹Calibration weights are applied to the data by the Home Office after the application of the design weights.

7.5 Analysis of non response

In 2009-10 the Home Office commissioned TNS-BMRB to carry out analysis of non response in the British Crime Survey.

The primary objectives of this analysis were to (a) explore the various components of non-response and what influences them, and (b) recommend a new weighting strategy based on the findings from (a).

Prior to 2009-10 the BCS has employed three weights: a design weight, a two-category non-response weight (the 'inner city' weight) and an additional 'calibration' weight to force certain distributions (sex, age and region) to match known population parameters.

The result of the non-response analysis was a recommendation to replace the inner city weight with a non-response weight generated by a model using the following predictive variables.

- Presence/absence of entry phone
- Whether/not in neighbourhood watch area
- Accommodation type
- Housing condition
- OAC
- PFA

This new three stage weight should incorporate more non-response information than the current weight and, consequently, may lead to a reduction in non-response bias.

The design effects of the different weights were compared and the design effect associated with the old inner city weight was 1.001; with the new non-response weight it was 1.007. Consequently, it was concluded that the new weight would reduce error. However, the final test of this will follow the calibration weight procedure.

Although a new set of 'shadow' weights for the 2009-10 survey have been computed as described above because the testing process was not yet completed it was decided to use the existing weighting schema as previous surveys for the 2009-10 results. Consequently any new weights will not be introduced until 2010-11 at the earliest. The new set of 'shadow' weights will be tested by the Home Office on the

victimisation rates and the final decision as to whether these new weights will be introduced will depend on the outcome of this testing process.

Full details of the non-response analysis can be found in Annex 1.

8. Comparing key survey variables with the population

The achieved sample was weighted in order to be representative of the population living in private households in England and Wales. A series of comparisons are presented in the following tables, showing to what extent the 2009-10 BCS achieved core sample reflected the population as a whole, after applying the appropriate design weights and before final calibration weighting.

The regional distribution of the adult population aged 16 years or over by Government Office Region compared with the Mid-2009 Population Estimates is shown in Table 8.1. This shows that the regional profile of the weighted sample was broadly in line with the population figures. The main discrepancy in the achieved sample was the under-representation of respondents in London compared with the population estimates. This reflects the lower response rates achieved in London as already noted.

Table 8.1 Comparison of the BCS core achieved sample compared with the population by Government Office Region, 2009-10 BCS

Government Office Region	Weighted Core Sample ¹	Mid-2009 population estimates ²²	Difference (sample – population)
	%	%	%
North East	5.1	4.8	+0.3
North West	13.0	12.6	+0.4
Yorkshire & The Humber	9.5	9.6	-0.1
East Midlands	7.5	8.2	-0.3
West Midlands	10.3	9.8	+0.5
East of England	10.6	10.5	+0.1
London	12.1	14.0	-1.9
South East	15.9	15.4	+0.5
South West	10.1	9.7	+0.4
Wales	5.5	5.5	0.0
<i>Bases:</i>	<i>45,189</i>	<i>44,554,400</i>	
<small>1 Prior to the calibration weights applied at a later stage by the Home Office. Source: Mid-2009 Population Estimates, Office of National Statistics</small>			

²² Adults aged 16 and over

Table 8.2 shows similar comparisons between the achieved core sample in relation to the Mid-2009 Population estimates for England and Wales by sex and age. This shows that the survey slightly under represented men, those aged under 35, and those aged over 85 (especially women). The profile of the survey by sex and age was very similar to the previous year. All of these patterns are fairly common in large scale surveys and reflect the slightly lower response rates achieved among these particular groups.

Although not reported here, as already mentioned the age and sex distribution of the achieved sample are further corrected by the Home Office at the analysis stage through the application of calibration weights so that the age and sex profile of survey respondents matched population estimates within each GOR (see [section 7.4](#)).

Table 8.2 Comparison of the BCS achieved core sample with the population by sex and age, 2009-10 BCS

	Weighted Core Sample¹	Mid-2009 population estimates	Difference (sample - population)
	%	%	%
Men			
16-19	6.0	6.7	-0.7
20-24	6.8	8.9	-2.1
25-34	13.1	16.6	-3.5
35-44	17.1	18.3	-1.2
45-54	17.9	16.7	+1.2
55-64	17.3	14.6	+2.7
65-74	13.2	10.3	+2.9
75-84	7.1	6.1	+1.0
85 and over	1.5	1.8	-0.3
<i>Bases:</i>	<i>20,360</i>	<i>21,731,200</i>	
Women			
16-19	5.7	6.1	-0.8
20-24	6.3	8.1	-2.2
25-34	13.7	15.4	-1.2
35-44	18.2	17.7	+1.4
45-54	17.9	16.3	+1.5
55-64	16.4	14.5	+1.2
65-74	12.6	10.7	+1.4
75-84	7.4	7.7	0.0
85 and over	1.9	3.7	-1.5
<i>Bases:</i>	<i>24,829</i>	<i>22,823,400</i>	
All men	47.1	48.8	-1.7
All women	52.9	51.2	+1.7
<i>Bases:</i>		<i>44,554,400</i>	
1 Prior to the calibration weights applied at a later stage by the Home Office. Source: Mid-2009 Population Estimates, Office of National Statistics			

9. The Under 16s survey

9.1 Background

The extension of the survey to include people under the age of 16 followed recommendations made by the Smith Review²³. The main rationale for extending the coverage of the BCS is to provide estimates of victimisation levels among under 16s so that their needs can be part of policy consideration and service delivery.

An extensive testing and development phase was undertaken prior to the launch of the survey which is described in detail in the development report²⁴.

In summary the development phases undertaken were:

- Qualitative research to explore young people's understanding of crime and victimisation
- Cognitive piloting of the questionnaire with young people
- Small scale test encompassing 20 interviewer points to test fieldwork procedures.
- Split sample experiment to assess the impact of the under 16s survey on the main survey
- Live survey launch in January 2009

In addition the Home Office conducted an external consultation to consult with stakeholders about the most appropriate age range to include in the survey.

9.2 10 to 15 year old sample

The methodological review commissioned by the Home Office recommended that the sample of children to be included in the survey should be obtained through identifying children in households already selected for the core BCS. The field procedures to be used would be broadly similar to those already used for the young adult boost, with screening being carried out after the core sample interview had been completed. The review also recommended that no more than one child was interviewed in each household.

²³ <http://www.homeoffice.gov.uk/rds/pdfs06/crime-statistics-independent-review-06.pdf>

²⁴ Extending the British Crime Survey to children: a report on the methodological and development work,

Identification of under 16s to take part in the survey was conducted in a very similar way to the procedures used to conduct the boost survey with 16 to 24 year olds. Screening was conducted at all core sample addresses where a productive outcome is achieved. Where an eligible child was identified one child was selected at random to take part in the interview. However in households with only one 10-15 year old present the child was only eligible to be interviewed in 87.5% of cases.

Adjusting the eligibility criteria in households where there was only one eligible child improved the precision of the estimates. The design effect due to base weight variance would be 1.18 when the number of interviews sought is independent of the number of eligible children in the household. It is closer to 1.05 when different rules are followed in households with only one eligible child. Ideally in households with only one eligible child an interview would have been attempted in 50% of cases but it was not possible to achieve the target number of interviews without raising this proportion to 87.5% of cases.

This screening process was managed through screening criteria included on the address contact sheet. Where only one 10-15 year old was identified interviewers were required to check whether the address contact sheet was labelled 'A' or 'B' and should they have an 'A' contact sheet no further interview would be conducted, Addresses were randomly allocated to the 'A' or 'B' classification prior to issue (87.5% allocated to 'B' and 12.5% allocated to 'A').

The aim was to achieve a sample of 4,000 interviews with 10-15 year olds for the 2009-10 survey.

9.3 Questionnaire content and development

An extensive development and testing phase was undertaken prior to the launch of the 10-15 survey, the results of which have been documented separately in the development report published in 2010.

The questionnaire for 10 to 15 year olds covered:

- Schooling and Perceptions of crime
- Crime screener questions – personal incidents only
- Victimization module
- Perceptions of and attitudes towards the police

- Anti-social behaviour
- Crime prevention and security
- Self completion module
 - Use of the internet
 - Personal safety
 - School Truancy
 - Bullying
 - Street gangs
 - Drinking behaviour
 - Cannabis use
 - Verification questions
- Demographics

Table 9.1 Modules of the 2009-10 Under 16s BCS questionnaire and sub-set of respondents who were asked each module

Questionnaire module	Proportion of sample	Module
Schooling and perceptions of crime	All	
Screeener questionnaire	All	
Victimisation modules	All victims	
Perceptions of and attitudes towards the police	Random 33%	A
Anti-social behaviour	Random 33%	B
Crime prevention and security	Random 33%	C
Use of the internet	Random 33%	C
Personal safety	All aged 13-15	
School truancy	All	
Bullying	Random 33%	B
Street gangs	Random 33%	A
Drinking behaviour	All	
Cannabis use	All	
Verification questions	All	
Demographics	All	

9.3.1 Schooling and perceptions of crime

This module includes questions about whether the respondent attends school and if so what school year they are in (school year is used later to help respondents recall exactly when incidents of crime took place)

A small number of questions are included about the respondent's perception of crime in their local area and whether they believe the level of crime has changed in the last few years.

9.3.2 Screener questionnaire

All respondents were asked whether they had experienced certain types of crimes or incidents within the last 12 months. To try and encourage accurate recall a life event calendar was offered to all respondents to act as a visual prompt when answering questions. 10 to 15 year olds were not asked about household incidents as these would have been covered in the interview with the adult household member. The 10-15 year olds were asked:

- Whether anything had been stolen from them
- Whether anyone had deliberately damaged their property
- Whether anyone had deliberately kicked, hit, pushed or been physically violent towards them in any other way
- Whether they had been hit or threatened with a weapon
- Whether they had been threatened in any other way

9.3.3 Victimization modules

All incidents identified at the screener questions are followed up in more detail in the victimisation module. Incidents are covered in specific priority order:

- According to the type of crime – As with the adult survey victimisation modules were asked in reverse order to the screener questions.
- Chronologically within each type of crime – If a respondent reported more than one type of incident of the same crime type, victim modules were asked about the most recent incident first and worked backwards chronologically
- Up to 6 mini victim forms completed with a maximum of four full victim forms

The 10-15 survey includes a mini victim form which is a short sub-set of questions used to identify low level incidents which were not subject to detailed questions. Respondents completed a mini victim form if the incident is classified as a 'relatively minor' incident. An incident is classified as relatively minor if all of the following apply:

- Incidents happened at school, where,
- The offender was a pupil at the victim's school, and,
- The offender did not use or threaten to use a weapon, and,
- The victim was not physically hurt in any way, and,
- Nothing was taken with the intention of taking it and not giving it back.

As with the core survey the victimisation module collects the key information required for classification of offences;

- The exact month in which the incident took place
- An open ended description of the incident
- Series of key questions to establish important characteristics of the incident

9.3.4 Perceptions of and attitudes towards the police

Respondents were asked their opinion about the police in their area and whether they agreed or disagreed with a number of statements about the police in the area. Questions were also asked about whether the respondent knew any police or community support officers (PCSOs), whether they had had any contact with police or PCSOs, who initiated the contact, reasons for contact and how satisfied they were with the contact.

9.3.5 Anti-social behaviour

This module included questions about whether respondents felt teenagers hanging around on the streets were a problem in the area, whether they themselves hung around on the streets and where they tended to hang around with friends. It also included questions about whether the respondent had been asked to move on from an area and if so by whom and for what reason. The module also included some questions about the types of activities available for young people in the area.

9.3.6 Crime prevention and security

Respondents were asked about when they go out in evening, whether they travel on public transport and if so how often and at what time of day. Questions were also included about whether they owned a mobile phone, MP3 player, games console or bike, and if so what precautions they took to protect these items.

9.3.7 Self completion modules

A number of modules contained particularly sensitive questions and were therefore included in the self completion section so that respondents did not have to tell the interviewer their answers. As in the core survey, practice questions were included so that the interviewer could explain to the respondent how to use the computer.

Use of the internet - respondents were asked whether they had used the internet in the last 12 months and if so what they used the internet for

Personal security – these questions were only asked of respondents aged 13 or over and covered whether they carried a personal alarm with them, whether the respondent knew anyone who carried a knife, whether they themselves carried a knife and if so why they did so.

School truancy – Three questions were asked covering whether the respondent had missed school without permission in the last 12 months, how many times they had missed school without permission and whether they had been suspended or excluded from school.

Bullying – This module asked whether the respondent had been bullied and where this was the case some follow up questions were asked about the nature and extent of the bullying.

Street gangs – This module included a definition of a street gang as;

Groups of young people who hang around together and do all of these things:

- *commit violent crimes together and ;*
- *Spend a lot of time in public places; and*
- *usually have a name, an area or territory, a leader, or rules*

Respondents were asked how much of a problem they believed street gangs to be in the country as a whole and in their local area. They were also asked whether they knew anyone who was a member of a street gang and whether they themselves were a member of a street gang.

Drinking behaviour – This module included questions about the respondent's perception of drinking behaviour among people their age, whether they had ever drunk alcohol, whether they had ever been drunk, and how often they had been drunk.

Cannabis use – Respondents were asked whether they had ever tried cannabis, how often they had tried it and whether they had felt stoned.

Verification questions – Two of the crime screener questions were repeated in the self completion section to explore whether respondents would give different answers if they did not have to say the answer out loud. The two screener questions included for verification were whether the respondent had been hit, kicked, pushed or assaulted in any other way and whether anyone had hit them or threatened them with a weapon. In addition respondents were asked whether they told the truth on all questions the interviewer asked them and whether they told the truth in the self completion section.

9.3.8 Demographics module

The demographics module included questions regarding whether the respondent had a disability.

9.4 Fieldwork procedures and documents

All respondents for the under 16s survey were selected from households already selected to take part in the core survey. Screening was only carried out in households where a successful adult interview was achieved. In most cases screening was conducted only on completion of the adult interview but in some cases screening was carried out before the adult interview had taken place.

Where a 10 to 15 year old was identified in a household interviewers were required to obtain the permission of a parent or guardian to interview the child before starting the

10 – 15 year old survey. Permission was recorded in writing on the address contact sheet. In some cases the adult respondent may not have been the parent or guardian of the child (for example an older sibling may have been interviewed in the core survey if they were aged 16 or over). In these cases interviewers were not able to obtain permission to interview the child from the core respondent and would therefore have to make contact with the parent or guardian to obtain permission.

Interviewers were provided with a parental information card which gave details of the nature and content of the survey and was to be presented to parents or guardians when they were asked for permission for the child to take part.

Once parental permission was obtained interviewers were instructed to ensure that the 10 to 15 year old also gave their consent to participate in the survey and that they understood what the survey would be about.

A leaflet was specifically designed for the 10 to 15 year olds that explained in relatively simple terms what the survey was about. Much of the detailed information about the survey was omitted from this leaflet on the basis that the 10 to 15 year olds would also have access to the original household letter and leaflet about the survey.

9.5 Item non-response

In order to emphasise to 10 to 15 year olds their right to refuse a particular question or the survey as a whole they were given a red and green card to use throughout the interview. If they chose not to answer a question they could simply present the interviewer with the red card and that particular question would be coded as a refusal.

The red and green card was developed primarily with the younger age groups in mind. It was also found useful however in reassuring parents that the 10 to 15 year olds could refuse certain questions if they felt uncomfortable.

9.6 Presence of others during the interview

The BCS interview for both adults and 10 to 15 year olds includes sensitive questions about whether the respondent has been a victim of crime or not. In general it is considered better to ask these types of questions when the respondent is alone as they will be more likely to answer the questions honestly and openly. In any

interview situation it is often not possible to control who else is present when the interview takes place and this was even more of an issue with the under 16s survey. Parents would often wish to be present for the interview or there may be siblings present at the time of interview.

The presence of others during the interview for the under 16s survey could be advantageous in some situations but detrimental in others. When a parent or guardian was present during the interview they could often help the 10 to 15 year old to feel more comfortable in the interview situation and could help to remind the child about incidents they may have forgotten about. From the point of view of interviewer protection it was also useful to have a parent/guardian present. However there was a concern that 10 to 15 year olds might not disclose certain incidents that had happened to them in front of others possibly because they were worried about getting into trouble or because they do not want to worry their parents.

As might be expected there was more likely to be someone else present during the screener questionnaire during the child interview than in the adult interview. During the adult interview there was no-one else in the room during 70% of interviews compared with just 28% of child interviews where there was no one else present (Table 9.1). Where there was someone else present during the child interview this was most likely to be a parent or guardian.

Parents or guardians were far more likely to be present during the interviews of younger children. In 83% of interviews with a 10 year old a parent or guardian was present compared with just over half (55%) of interviews with 15 year olds. Parents or guardians were slightly more likely to be present during the screener interview for female respondents compared with male respondents (71% and 67% respectively).

**Table 9.1 Presence of others during the screener questionnaire, 2009-10
Under 16s survey**

	Age of child						Total
	10	11	12	13	14	15	
	%	%	%	%	%	%	%
Parent/guardian	83	76	74	67	61	55	69
Other child from household	12	11	10	9	8	7	9
Other adult from household	1	2	1	4	2	3	2
Other non-household child	2	3	2	2	3	3	2
Other non-household adult	2	1	1	1	2	2	2
No one present	13	21	24	29	36	40	28
<i>Base:</i>	<i>549</i>	<i>621</i>	<i>667</i>	<i>641</i>	<i>675</i>	<i>650</i>	<i>3,803</i>

9.7 Acceptance of self completion and use of Audio-CASI

Almost all 10-15 year olds (99%) completed the self-completion section of the survey compared with 93% of adult respondents. Most (97%) completed the self completion section themselves, two per cent asked the interviewer to complete it for them and just one per cent refused to complete it.

Ten year olds were slightly less likely than those aged 11 or older to complete the section themselves (94% compared with 98% of children aged 11 or older).

The self-completion section was offered to respondents with an option to use audio-CASI. Audio-CASI enables the respondent to listen to the questions being read out through a set of headphones. In this way the privacy of the interview is maintained whilst allowing the respondent to understand the questions if they have difficulty reading them themselves. Table 9.2 shows the proportion of respondents who used the audio-CASI.

A third (33%) of respondents used audio CASI for all or most questions while 64% did not use it at all. Audio CASI was more likely to be used by younger respondents

(40% of 10 year olds used it for all or most questions compared with 27% of 15 year olds).

Table 9.2 Use of Audio-CASI for self completion by age of respondent

	Age of child						Total
	10	11	12	13	14	15	
	%	%	%	%	%	%	%
Used for all or most questions	40	35	34	33	32	27	33
Used for some/few questions only	4	4	3	3	3	3	3
Not used	57	61	64	65	65	70	64
<i>Base:</i>	<i>541</i>	<i>614</i>	<i>664</i>	<i>635</i>	<i>670</i>	<i>647</i>	<i>3,771</i>

9.8 Interview Length

As with the core survey timing stamps were present throughout the under 16s survey questionnaire to measure the interview length. Some unusually short or long interview times were removed (see section 4.6.1 for an explanation of outlying interview times). Any times of less than 8 mins or more than 120 were removed. For the under 16s survey 98% of interviews had a valid time and are included in the analysis that follows.

Table 9.3 shows the average interview length by type of respondent. The average (mean) interview length of the under 16s survey during 2009-10 was 20 minutes. Interviews tended to be slightly longer on average for males and for younger respondents. As is the case with the core interview respondents who report having been a victim of at least one crime have a longer average interview length, 29 minutes compared with 16 minutes for non-victims.

Table 9.3 Average interview length by sex, age and victimisation

	Average time (minutes)	Base
Sex		
Male	21	1,871
Female	19	1,805
Age		
10	23	513
11	22	603
12	20	657
13	21	619
14	19	659
15	19	625
Victimisation		
Victim	29	2,463
No victim	16	1,213

9.9 Response rate and reasons for non response: 10-15 year old sample

Table 9.4 shows the screening and response outcomes for the 10-15 year old sample. During 2009-10, interviewers were required to screen for 10 to 15 year olds at all of their core sampled addresses where a core interview was conducted. After accounting for deadwood addresses, 24% of addresses which were issued for the core survey were not screened for 10-15 year olds because the outcome at the core address was an unsuccessful outcome.. Interviewers identified at least one 10-15 year old at 13% of addresses where screening was successfully carried out. Among those households where an eligible respondent was identified the response rate achieved was 68%.

The level of non-contact (2%) was broadly in line with the level achieved on the core sample but the levels of refusals were slightly higher at 21%. The response rate achieved on the 10 to 15 year olds survey does not take into account households where it was not known whether a 10-15 year old was present because of non-

response to the core sample. When this is taken into consideration the 'true' response rate for the 10-15 survey is 52%²⁵.

²⁵ This is calculated by applying the actual eligibility rate achieved for successfully screened addresses (13.3%) to the total non-deadwood addresses issued for screening (59742) to give an estimate of 7,346 eligible households, from which 3,803 interviews were achieved which represents a response rate of 52%.

Table 9.4 Screening outcomes and response outcomes for the 10-15 year old sample, 2009-10 BCS

	N	% of issued eligible addresses	% of screened households	% of eligible households
TOTAL ADDRESSES FOR SCREENING	66,861	100.0		
<i>Core deadwood addresses</i>	7119			
TOTAL ELIGIBLE ADDRESSES FOR SCREENING	59,742	100.0		
No screening attempted (eligibility unknown)	14,553	24.4		
Screening information refused (eligibility unknown)	24	0.0		
Total unknown eligibility	14,577	24.4		
Total households screened for 10-15 year olds	45,165	75.6	100.0	
Screened households with no a 10-15 year old	39,139	65.5	86.7	
Screened households with a 10-15 year old	6,026	10.1	13.3	
Total screened households with a 10-15 year old	6,026		100.0	
10-15 year old in household, no interview required	468		7.8	
10-15 year old in household, interview required	5,558		92.2	
Total households where interview required	5,558			100.0
No contact with selected respondent	106			1.9
No contact with parent/guardian	18			0.3
Total non-contact	124			2.2
Office refusal	2			0.0
Parent/guardian permission refusal	708			12.7
Personal refusal	354			6.4
Proxy refusal	57			1.0
Contact made, no specific appointment	31			0.6
Total refusal	1,152			20.7
Other unproductive				
Broken appointment	39			0.7
Temporarily ill/incapacitated	8			0.1
Physically or mentally unable	73			1.3
Away/in hospital	28			0.5
Inadequate English	6			0.1
Other unsuccessful	325			5.8
Total other unsuccessful	479			8.6
TOTAL UNPRODUCTIVE	1,755	2.6		31.6
Full interviews	3,803			68.4
Partial interviews	0			0.0
TOTAL INTERVIEWS	3,803			68.4

9.10 Data processing

9.10.1 Offence coding

The offence coding system used for the under 16 survey was based on the system designed for the core survey but was adapted to be suitable for the types of incidents experienced by 10 to 15 year olds. Full details of the development of the coding system can be found in the Development report.

The main difference between the core and under 16s coding is the use of summary offence codes. These a codes are applied to low level incidents where only summary information is collected through the mini victim form rather than full and detailed information that is collected in the full victim form.

These summary codes are:

Code 116 – Aggressive behaviour

Code 117 – Attempted aggressive behaviour

Code 146 – Theft with aggressive behaviour

Code 147 – Attempted theft with aggressive behaviour

Code 166 – Theft

Code 167 – Attempted theft

Code 186 – Damage to property

Code 187 – Attempted damage to property

Code 196 – Intimidation/coercion

9.11 Home office coding

As with the core survey all cases which the coders are uncertain about are referred to the Home Office for further verification. In addition 20% of all codes which TNS-BMRB were certain about were selected and sent to the Home Office for quality control checking. This is a higher proportion of cases than is sent for the core survey which reflects the fact that the offence coding system has recently been developed and requires additional quality checks to ensure all scenarios have been covered in the guidance. The results from the 2009-10 under 16s survey were published as Experimental Statistics reflecting the fact that the classification of offences was still in development.

In total 459 cases were sent to the Home Office for checking as part of the 2009-10 under 16s survey, which represented around 25% of all victimisation modules.

Of the victimisation modules sent to the Home Office:

- 15 were automatically referred to the Home Office. This covers cases including any sexual element, duplicate cases and cases where the victimisation module was invalid.
- 71 cases where the TNS-BMRB coder was not certain about the code
- 314 were part of the quality control check
- 59 were related victimisation modules

Of the 459 victimisation modules referred to the Home Office 88 had their code changed by the Home Office, representing 19% of all cases sent. This is a higher percentage than for the core survey and reflects the fact that the coding system for the under 16s survey had been newly developed and queries were still being raised across the coding teams as to the accurate code to be applied in certain cases.

The codes changed by the Home Office according to the categories outlined were as follows:

- In 10 cases offences were coded 01 for referral to the Home Office. As code 01 is not a valid code this code was changed in all cases
- In 5 cases referred as duplicates 2 were changed
- Of the 71 cases where TNS-BMRB coders were uncertain 38 (54%) were changed.
- Of 314 cases sent as part of the quality control check 38 had their codes changed (12%)
- Of the 59 related forms none had their codes changed.

In all cases where the Home Office changed a code the code was reviewed by the TNS-BMRB coders. In total 18 cases were referred back to the Home Office with queries regarding the change made and in 6 cases the original TNS-BMRB code was restored. After all queries had been resolved 82 cases were changed by the Home Office, representing 18% of all cases sent.

9.11.1 Final offence code

The SPSS set delivered to the Home Office includes all the offence codes that have been given to each victimisation Module at every stage of the coding process. It also includes an additional variable 'Offclass' which defines whether an incident is classified as a 'relatively minor' incident or as a 'relatively serious' incident. This classification is not part of the coding process but is derived in SPSS based on answers to a small set of questions coded by the coders covering:

- Whether there was INTENTION to steal, hurt or damage
- Whether the victim knew the offender
- The level of any hurt inflicted or cost of items stole or damaged

The same consistency checks as are run on the adult data are run on the under 16s data to check the offence code. In addition checks are run to ensure that any serious offence codes (such as wounding etc) have not been classified as relatively minor offences and that summary or full offence codes have been applied correctly to mini and full victim forms respectively.

9.12 Data output

The data for the under 16s survey is delivered to the Home Office to the same quarterly timetable as the core survey data. As with the core data two data files are supplied, the Non Victim File and the Victim File.

Due to the development of the offence classification system the first full data file including the offence code for the under 16s survey was delivered to the Home Office in March 2009. Since that initial delivery data has been delivered to the core survey delivery timetable.

9.13 Weighting

In addition to exploring non-response on the adult survey (see [section 7.5](#)) a non-response weight was calculated for the 10-15 year old survey.

The variables that were found to be significantly associated with non-response were included in the final model. This model used logistic regression to obtain the probability of response based on the following variables:

- whether sampled child had mobile phone stolen (no phone; has phone-not stolen; has phone-stolen)
- length of adult interview (banded <1h30, 1h30+)
- Main newspaper readership (broadsheet, Tabloid, other/no main paper, none)
- Whether Adult accepted self completion (Yes, No)
- How confident are you that the police are effective at catching criminals
- Number of adults in the household (1,2,3,4,5+)
- Age of child sampled

The following were not significant, but were included for completeness:

- Whether adult is a victim of crime
- Sex of sampled child.

Creating the final weights

There were several steps to creating the final weight for the under 16s survey. The non-response weight that incorporates the design weight for the number of eligible children in the household is based on responding households. The household non-response weight from the core adult file is multiplied by the child non-response weight to give an overall unscaled and untrimmed child weight. This was capped at the 99th percentile so as to reduce the impact of any unusual, large weights, and then scaled so that the weighted sample size matched that of the achieved sample size. Full details of the non response analysis can be found in Annex 2.

Annex 1 – Non-response analysis (based on 2008-09 survey)

Background

In 2009-10 the Home Office commissioned TNS-BMRB to carry out analysis of non response in the British Crime Survey.

The primary objectives of this analysis were to (a) explore the various components of non-response and what influences them, and (b) recommend a new weighting strategy based on the findings from (a).

This chapter describes a model of case-level response propensities derived from auxiliary data available about each sampled case (address) issued for the **2008-09** survey.

Much of this auxiliary data is drawn from the census although it is only available aggregated at a small area level. For the most part TNS-BMRB has favoured aggregates at the LSOA level because, although it is possible to use the smallest aggregation level (Census Output Area or 'OA'), LSOAs are less likely to have changed profile since 2001.

There are also some important higher level geographies – e.g. council wards or Police Force Areas – but smaller aggregations have been prioritised over larger ones. Nevertheless there is residual predictive power associated with these administrative identities.

A large number of area-based variables *could* be included in this analysis but TNS-BMRB has restricted the models to the current range of geo-demographic data appended to the dataset. This is a sufficiently deep pool to choose from.

As well as area-based data, the vast majority of sampled cases also have 'interviewer observations' about the state of the sampled address itself and of the surrounding streets/neighbourhood. Because this data is specific to each sampled case, it generally takes priority over the aggregated area level data.

It is worth noting here that the following analysis is based only on cases with all of these 'interviewer observations' available. As one of the purposes of this study is to generate new non-response weights, this presents a problem. Our suggestion is that

a regression procedure is used to impute weights for those with missing values on one or more of the interviewer observations²⁶.

Summary of results

TNS-BMRB employed a series of nested binary logistic regression models to explore non-response, treating it as the result of a three-stage decision process:

1. Contact with sampled household? (Y/N)
2. (If Y) Achieved interview? (Y/N)
3. (if N) Achieved interview at re-issue stage? (Y/N)

Household contact

The model of household contact is the strongest model of the three but still fairly weak by any standards. Only four variables are significant positive predictors of contact:

- no entry phone
- detached or semi-detached house
- fairly or very good housing condition
- more settled areas, particularly rural or semi-rural locations or those with lots of older or 'blue collar' residents

Entry phones both prevent direct contact between interviewers and sampled households and are indicative of households that guard their privacy. Entry phones are most commonly found in flats but the model shows that both accommodation type and entry phone status have additional independent predictive power.

Poor housing condition might be indicative of housing that is not frequently occupied but there is no particular evidence for that among the respondent group.

The more urban segments - with significant population churn and long working hours - are associated with relatively low contact rates while more settled areas - particularly if in rural or semi-rural locations - are associated with relatively high contact rates.

²⁶ 4699/45519 interview cases 10.3% were missing at least one of the observations used in the final model (usually whether or not address was in a neighbourhood watch scheme).

Co-operation

The model of co-operation (*given* household contact) was the weakest model of the three. This suggests that the address-level details collected by the interviewers - while of some use predicting contact rates - are not particularly useful for predicting co-operation rates.

Nevertheless, five variables are significant predictors of co-operation:

- Neighbourhood Watch area
- no entry phone,
- fairly or very good housing condition
- self-contained property
- areas with higher than average contact probabilities excluding London

Several of these variables are also significant predictors of household contact, reflecting the well-known correlation between contact rates and co-operation rates.

Neighbourhood Watch areas tend to be found in more settled, safer areas in which there may be less suspicion about household callers than there is elsewhere. This might sound paradoxical since a Neighbourhood Watch scheme is indicative of a community concerned with security. However, it is more strongly indicative of a community that wants to keep the generally high level of security they enjoy.

Entry phones not only make contact difficult but are indicative of households that guard their privacy and purposefully minimise contact with the world outside their doors.

The association with accommodation type is much less strong than is the case with contact. Some types of flat are associated with a fairly *high* level of co-operation once entry phone status is controlled for.

However, property that is not self-contained is strongly indicative of non-cooperation just as it is strongly indicative of non-contact.

The association between housing condition and co-operation rates is not obviously explicable although a lack of care with regards to the building's upkeep may reflect a lack of interest in the surrounding community and, perhaps, a lack of interest in

surveys like the BCS which are (partially) sold on the basis that they have a strong social value.

The types of area where contact probabilities are high are, for the most part, the types of area where co-operation probabilities are high. There are exceptions. London-specific areas tend to have lower than expected co-operation rates.

Re-issue success

This last model covers both contact and co-operation rates across multiple re-issue stages. Cases may be re-issued up to four times although only a tiny fraction are, given the need to complete fieldwork assignments within a four month timeframe.

Re-issues also stretch fieldwork resources in a geographically uneven fashion which means that unproductive cases may be more or less likely to be re-issued depending on local resource levels.

Another reason why multiple re-issues is rare is the low productivity rate. The most productive re-issues are non-contacts. Controlling for the other variables in the model, the probability of achieving an interview from an original household non-contact is twice that of an original refusal. However, the probability is still low (18% compared to 9%).

The only location-based variable that proves significant is Police Force Area, and this is probably due to the local variation in re-issue rate rather than anything else.

Model development

Introduction

TNS-BMRB employed a series of nested binary logistic regression models to explore non-response, treating it as the result of a three-stage decision process²⁷:

1. Contact with sampled household? (Y/N)
2. (If Y) Achieved interview? (Y/N)
3. (if N) Achieved interview at re-issue stage? (Y/N)

There are various indicators of the explanatory power of logistic regression models. TNS-BMRB elected to use the Nagelkerke 'pseudo R²' which gives an estimate of the amount of variance in a binomial dependent variable that is explained by the predictor variable (or group of variables) entered into the model. Table 10.1 shows this value for each of the three stages above.

Table 10.1: Model power (3 stage response process)

Model	Variance explanation (NR ²)
1. Contact with sampled household? (Y/N)	9.7%
2. (If Y) Achieved interview? (Y/N)	2.2%
3. (if N) Achieved interview at re-issue stage (Y/N)	6.4%

It is immediately clear from this table that (a) all of the models are relatively weak, and (b) the auxiliary data mainly helps to explain household contact variance rather than co-operation variance.

The model for stage 3 is somewhat stronger than that for stage 2 because the outcome of stages 1 and 2 can be included in the model for stage 3. However, it is still a fairly weak model.

²⁷ Draft models were based on a six-stage decision process but BMRB found that the auxiliary data produced very weak models for response stages 2-4. While this is informative – helping identify where differential non-response is located within the general response process – the details of each model are largely irrelevant. Consequently, BMRB decided to simplify the response process into three stages.

Model building methods

Most of the work was carried out using the logistic regression package in SPSS 16.0. The general method of working was to enter each group of variables in ascending geographical order, so the first group entered was the address-level group, the second group entered was the OA-level group, and so on.

A 'stepwise' procedure was used to identify the smallest number of significant predictors in each group. The intention was to retain the identified predictors in the model, even if they lost statistical significance due to the addition of a subsequent group of variables. In practice, this issue did not arise.

Care was taken to avoid including variables that overlap too much (i.e. explain the same component of the variance) but a degree of overlap is inevitable and is noted in the text.

The Hosmer-Lemeshow test was used to assess model fit. This is a chi-square test in which cases are divided into deciles based on modelled response propensity. The expected number of responding cases in each class is then compared with the actual number of responding cases. A significant test result suggests that the model works better for one end of the distribution than the other (a phenomenon called 'heteroscedasticity').

All Hosmer-Lemeshow tests were non-significant suggesting that heteroscedasticity is not a problem for any of these models. When it *is* a problem, the normal solution is to include some interaction terms in the model but these were not required on this basis and, in any cases, none were found to be statistically significant.

The final estimates were computed using the 'complex samples' logistic regression package in SPSS 16.0. This is slightly less flexible than the basic package – which is why it was not used for the model development stage – but is vital for computing the correct confidence intervals around each parameter estimate.

The parameter estimates for each model are detailed in the following sections.

Selection of auxiliary variables

The first stage of this project was to review the collection of auxiliary variables, rationalise the code-frames where necessary and select representative variables where several were highly inter-correlated.

With one or two exceptions, these variables are categorical. The census variables in their raw forms are quantitative but were also converted into *deciles*, an (ordinal) categorical form.

The auxiliary variables can be divided into five types on the basis of relevant geography. These types and associated variables are shown in table 10.2 below.

Table 10.2: Auxiliary variables by type.

Variable type	Variables [variable names]
Address-level	<p>Interviewer observations:</p> <ul style="list-style-type: none"> • Type of dwelling [acctyp2] • Visible security features [vissec2a-h] • Location of dwelling within neighbourhood [locatea-g] • Housing condition relative to neighbourhood [relcond2] • General state of housing in neighbourhood [houcond2] • Prevalence of poor housing, vandalism and litter in neighbourhood [poorhou2, vandcomm2, rubbcomm2] • Whether in a neighbourhood watch area [neigwat2]
OA level	<p>Small area segmentations:</p> <ul style="list-style-type: none"> • Output Area Classification (OAC) with extra London-only categories [oapfa] • <i>A Classification of Residential Neighbourhoods (ACORN) [agr2001 and other levels]</i> • ONS urban-rural scale [rural]
LSOA level	<p>Census data:</p> <ul style="list-style-type: none"> • <i>Proportion of population that is economically active</i> • Proportion of economically active population that is unemployed • <i>Proportion of population aged 16-64 that are managers or professionals</i> • Proportion of population that is retired • Proportion of population that is a student • <i>Proportion of population that is not White</i> • Proportion of population that is Asian • Proportion of population that is Black
MSOA level	Clustering method used in sample design [stratum]
Other larger geographies	<ul style="list-style-type: none"> • <i>ONS Ward classification [wrddgrp]</i> • <i>Ward level indices of deprivation (separate indices in England and in Wales)</i> • Police Force Area [pfacode] (model 3 only)

This initial list was further edited after inspection of the data. Where variables were highly inter-correlated only one was included in the models. The omitted variables are shown in italics in Table 2. The key choice between Output Area Classification and Police Force Area is detailed below.

Police Force Area and the Output Area Classification

Police Force Area is the primary stratification variable used for the BCS and ensures a sample that is regionally representative. It was a powerful predictive variable in draft versions of all three models but presented a problem due to its relatively strong correlation with the Output Area Classification (OAC).

For example, 'Afro-Caribbean communities' are almost all in London. The household contact rate is fairly low in these areas but the early models 'loaded' this low contact probability on to the OAC identity rather than to the PFA identity. The same thing happened with regard to several other London-centric OAC categories ('Transient Communities', 'Settled in the City' and 'Asian Communities'). Consequently, London had a relatively high contact probability *controlling for OAC*. However, this does not make much sense since what makes London London is the nature of its component localities.

There are only two solutions to this problem: (1) create a combined PFA-OAC variable or add an interaction term to the model, or (2) exclude one or the other of the two variables.

A version of the first option has been selected for the first two models (covering the probability of household contact and the probability of interview given household contact). This variable is based on the OAC which, despite having less predictive power than PFA identity, is undeniably more informative about the nature of the neighbourhood. The four London-centric categories ('Transient Communities', 'Settled in the City', 'Asian Communities' and 'Afro-Caribbean communities') have been split into London/Elsewhere. More detailed breakdowns were considered but they tended to result in too many cells with a small number of cases and did not improve the overall fit of the models.

PFA was retained in its entirety for the third model (reissue interview probability) because (a) the main effect of OAC was not statistically significant, and (b) cases that are eligible for re-issue are not always re-issued; there is a strong geographic bias best represented via the PFA variable.

Model 1: Propensity for household contact

At-a-glance

Dependent variable	Whether contact made with at least one adult in the household at first issue stage
Excluded cases	Any coded at first issue stage as un-issued ²⁸ ('office refusals') or as ineligible ('deadwood')
Significant positive predictors	<ul style="list-style-type: none"> • No entry phone • Detached or semi-detached whole house • Fairly or very good housing condition • Particular OAC segments (top 3 = 'prospering older families', 'village life' and 'older blue collar')
Significant negative predictors	<ul style="list-style-type: none"> • Entry phone • Flat, especially if not self-contained • Middling to very bad housing condition • Particular OAC segments, especially inner city ones (bottom 3 = 'transient communities' in and outside of London, 'Afro-Caribbean communities' outside of London)
Nagelkerke R ²	9.7%

Summary

This was the strongest model of the three yet only accounts for 9.7% of the variance in household contact levels, suggesting that the address-level details collected by the interviewers are no match for more specific details about the individuals who make up the household.

Only three of the address-level variables are significant predictors of contact: the presence or otherwise of an entry phone, accommodation type and the condition of the housing.

The direction of influence for these three variables is as expected.

Entry phones both prevent direct contact between interviewers and sampled households and are indicative of households that guard their privacy. Entry phones

²⁸ It is possible to model the propensity to deliver an office refusal but the lack of address-level data for most (but not all) cases means that this model is weak ($NR^2 = .01$). Only one predictor – OAC group – proved significant. Of course, it could be argued that an office refusal *is* contact of a sort. However, they have not been included as contacts here because the process of contact is clearly quite different from that of doorstep contact.

are most commonly found in flats but the model shows that both accommodation type and entry phone status have additional independent predictive power.

It is worth illustrating this by computing some exemplar contact probabilities for two different housing types with/without entry phone and keeping other variables in the model at their reference categories (fairly good housing condition, 'settled households' OAC segment):

Semi-detached house without entry phone:	97.5%
Semi-detached house with entry phone:	96.0%
Self-contained conversion flat without entry phone:	91.1%
Self-contained conversion flat with entry phone:	86.4%

Poor housing condition was also associated with non-contact and might be indicative of housing that is not frequently occupied. However, there is no particular evidence for that among the respondent group. It seems more likely that some of these cases may be 'deadwood' that is not identified as such. If so, then the influence of this variable in the model is somewhat overstated.

Contact rates by OAC segment make sense. The more urban segments - with significant population churn and long working hours - are associated with relatively low contact rates while more settled areas - particularly if in rural or semi-rural locations - are associated with relatively high contact rates. The high contact rates associated with areas containing 'blue collar' workers or older people is particularly notable, perhaps reflecting that fact that these areas contain many families with one or no earners which means that at least one person is likely to be in when an interviewer calls (daytime and early evening).

Table 10.3 shows the base odds of household contact as well as odds ratios associated with each category of each predictor. Both the base odds and the odds ratios are shown in the column labelled 'Exp(B)'. The next two columns show the 95% confidence intervals for these statistics. Large intervals generally reflect small sample sizes while small intervals generally reflect large sample sizes.

The overall odds for a particular combination of categories is computed by taking the base odds and multiplying it by the product of the category odds ratios. The predicted contact probability follows via this equation: $\text{probability} = \text{odds} / (\text{odds} + 1)$.

For example, the odds of contact for a self-contained purpose built flat with an entry phone, in very good condition and located in the 'thriving suburbs' is $39.38 \times 0.32 \times 0.62 \times 1.04 \times 0.93 = 7.56$. The contact probability follows as $7.56/8.56 = 88.3\%$.

Table 10.3: Odds ratios for household contact

Parameter	Exp(B)	95% Confidence Interval for Exp(B)	
		Lower	Upper
BASE ODDS	39.38	31.27	49.59
ACCOMMODATION TYPE			
Non-self-contained conversion flat	0.06	0.04	0.09
Non-self-contained purpose built flat	0.15	0.10	0.23
Bedsit / rooms	0.22	0.10	0.47
Self-contained conversion flat	0.26	0.19	0.35
Self-contained purpose built flat	0.32	0.26	0.40
Maisonette	0.51	0.34	0.78
Mid-terrace whole house	0.55	0.46	0.64
Caravan / mobile home	0.55	0.26	1.19
End of terrace whole house	0.61	0.51	0.73
<i>Semi-detached whole house</i>	<i>1.00</i>	.	.
Detached whole house	1.12	0.92	1.36
WHETHER HAS ENTRY PHONE			
Yes	0.62	0.53	0.74
<i>No</i>	<i>1.00</i>	.	.
CONDITION OF ACCOMMODATION			
Very bad	0.58	0.35	0.95
Fairly bad	0.50	0.41	0.60
Neither good nor bad	0.77	0.67	0.88
<i>Fairly good</i>	<i>1.00</i>	.	.
Very good	1.04	0.92	1.18
OAC GROUP / LONDON			
Transient Communities –London	0.46	0.32	0.66
Afro-Caribbean Communities – not London	0.49	0.31	0.77
Transient Communities – not London	0.61	0.39	0.96
Settled in the City – not London	0.88	0.64	1.22
Public Housing	0.88	0.60	1.29
Afro-Caribbean Communities – London	0.90	0.60	1.34
Thriving Suburbs	0.93	0.67	1.28
Young Families in Terraced Homes	0.93	0.70	1.22
Prospering Younger Families	0.94	0.65	1.36
Asian Communities – not London	0.95	0.69	1.30
<i>Settled Households</i>	<i>1.00</i>	.	.

Aspiring Households	1.02	0.76	1.39
Senior Communities	1.04	0.67	1.62
Asian Communities – London	1.04	0.75	1.44
Least Divergent	1.10	0.83	1.45
Younger Blue Collar	1.12	0.84	1.50
Settled in the City – London	1.13	0.72	1.78
Accessible Countryside	1.14	0.80	1.62
Prospering Semis	1.18	0.87	1.59
Older Workers	1.22	0.93	1.60
Agricultural	1.25	0.82	1.92
Terraced Blue Collar	1.33	0.91	1.96
Older Blue Collar	1.36	1.00	1.84
Village Life	1.38	0.99	1.92
Prospering Older Families	1.53	1.02	2.30

Model 2: Propensity for interview, given household contact

At-a-glance

Dependent variable	Whether interview achieved at first issue stage, given contact made with at least one adult in the household
Excluded cases	Any coded at first issue stage as un-issued ('office refusals') as ineligible ('deadwood') or where no contact made with household
Significant positive predictors	<ul style="list-style-type: none"> • Neighbourhood Watch area • No entry phone • Fairly or very good housing condition • Self-contained property • Particular OAC segments, especially rural ones (top 3 = 'agricultural', 'village life' and 'accessible countryside')
Significant negative predictors	<ul style="list-style-type: none"> • Not a Neighbourhood Watch area • Entry phone • Middling to bad housing condition • Non-self-contained property • Particular OAC segments, especially immigrant London ones (bottom 3 = 'settled in the city' (London) 'Asian communities' (London) and 'Afro-Caribbean communities' (London))
Nagelkerke R ²	2.2%

Summary

This was the weakest model of the three, accounting for only 2.2% of the variance in co-operation levels. Models as weak as this would not normally be published and it suggests that the address-level details collected by the interviewers - while of some use predicting contact rates - are not particularly useful for predicting co-operation rates.

Nevertheless, four of the address-level variables are significant predictors of co-operation²⁹: whether located in a Neighbourhood Watch area, the presence or otherwise of an entry phone, accommodation type and the condition of the housing. Three of these were also significant predictors of household contact, reflecting the well-known correlation between contact rates and co-operation rates.

The direction of influence for these four variables is largely as expected.

²⁹ In fact, a few more variables reached statistical significance (presence of burglar alarm or security gate plus location on a housing estate were all negatively correlated with co-operation) but they added so little to the model that they were excluded for reasons of clarity.

Neighbourhood Watch areas tend to be found in more settled, safer areas in which there is less suspicion about household callers than there is elsewhere. This might sound paradoxical since a Neighbourhood Watch scheme is indicative of a community concerned with security. However, it is more strongly indicative of a community that wants to keep the generally high level of security they enjoy and that allows them the freedom to 'trust' the intentions of household callers.

Entry phones not only make contact difficult but are indicative of households that guard their privacy and purposefully minimise contact with the world outside their doors. The contact model also found that accommodation type had additional predictive power with those most likely to have entry phones (flats) presenting an intrinsic barrier to contact, regardless of whether there is an entry phone or not. In this model, the association is much less strong. Some types of flat are associated with a fairly *high* level of co-operation once entry phone status is controlled for.

However, property that is not self-contained is strongly indicative of non-cooperation just as it is strongly indicative of non-contact. If we set other variables at their reference categories the total probability of achieving an interview (i.e. taking account of both contact and co-operation propensity) at a semi-detached house as 76.3%, while the probability of achieving the same at a non-self-contained purpose built flat is 49.9%. Despite the weakness of the two models, that is a significant difference.

Poor housing condition was associated with non-cooperation as well as non-contact although the odds of co-operation are unusually high when the housing condition is described as 'very bad'. This last may be due to the small sample size in this category rather than reflecting an interesting u-shaped probability distribution.

The OAC segments line up in similar fashion as with model 1 (contact probability). In fact, the correlation between the ranks is 0.63. The only distinctive difference is that those in the London 'settled in the city' and 'Asian communities' areas are less likely to co-operate than one might expect, given their contact probabilities. The reverse is true of those in the 'public housing', 'younger families in terraces' and 'settled in the city – not London' areas.

Table 10.4 shows the base odds of household contact as well as odds ratios associated with each category of each predictor. Both the base odds and the odds ratios are shown in the column labelled 'Exp(B)'. The next two columns show the 95% confidence intervals for these statistics. Large intervals generally reflect small sample sizes while small intervals generally reflect large sample sizes.

The overall odds for a particular combination of categories is computed by taking the base odds and multiplying it by the product of the category odds ratios. The predicted contact probability follows via this equation: probability = odds / (odds+1).

Table 10.4: Odds ratios for co-operation

Parameter	Exp(B)	95% Confidence Interval for Exp(B)	
		Lower	Upper
BASE ODDS	3.60	3.21	4.04
ACCOMMODATION TYPE			
Non-self-contained purpose built flat	0.39	0.28	0.56
Non-self-contained conversion flat	0.43	0.25	0.76
Maisonette	0.93	0.77	1.13
End of terrace whole house	0.94	0.87	1.03
Mid-terrace whole house	0.97	0.90	1.05
Self-contained conversion flat	0.97	0.84	1.13
<i>Semi-detached whole house</i>	<i>1.00</i>	.	.
Bedsit / rooms	1.02	0.58	1.81
Detached whole house	1.05	0.97	1.13
Self-contained purpose built flat	1.17	1.01	1.35
Caravan / mobile home	1.58	0.96	2.60
WHETHER HAS ENTRY PHONE			
Yes	0.64	0.56	0.74
<i>No</i>	<i>1.00</i>	.	.
CONDITION OF ACCOMMODATION			
Very bad	1.13	0.77	1.67
Fairly bad	0.71	0.63	0.80
Neither good nor bad	0.85	0.80	0.90
<i>Fairly good</i>	<i>1.00</i>	.	.
Very good	1.15	1.09	1.22
WHETHER VISIBLY IN N'HOOD WATCH AREA			
<i>No</i>	<i>1.00</i>	.	.
Yes	1.24	1.15	1.33
OAC GROUP / LONDON			
Settled in the City – London	0.66	0.56	0.79
Asian Communities – London	0.70	0.60	0.82
Afro-Caribbean Communities – London	0.72	0.62	0.84
Transient Communities – not London	0.74	0.54	1.01

Transient Communities – London	0.75	0.58	0.97
Asian Communities – not London	0.84	0.70	1.01
Afro-Caribbean Communities – not London	0.91	0.62	1.35
Prospering Younger Families	0.93	0.79	1.11
Senior Communities	0.93	0.73	1.19
Thriving Suburbs	0.97	0.83	1.13
Younger Blue Collar	0.98	0.85	1.12
Settled in the City – not London	0.99	0.82	1.20
Aspiring Households	0.99	0.86	1.15
<i>Settled Households</i>	<i>1.00</i>	.	.
Young Families in Terraced Homes	1.03	0.87	1.21
Older Workers	1.04	0.90	1.20
Public Housing	1.06	0.85	1.33
Least Divergent	1.06	0.92	1.22
Older Blue Collar	1.07	0.93	1.24
Prospering Older Families	1.10	0.95	1.27
Prospering Semis	1.12	0.97	1.29
Terraced Blue Collar	1.15	0.96	1.37
Accessible Countryside	1.16	0.99	1.37
Village Life	1.23	1.06	1.42
Agricultural	1.24	1.02	1.50

Model 3: Propensity for re-issue interview

At-a-glance

Dependent variable	Whether interview achieved at re-issue stage, given no interview at the first issue stage
Excluded cases	Any coded at any issue stage as un-issued ('office refusals') as ineligible ('deadwood') or at the first stage as interview
Significant positive predictors	<ul style="list-style-type: none"> • Non-contact at the first issue stage • Particular Police Force Areas (top 3 = Durham, Dyfed Powys and Northumbria) • Reasonably high first issue contact probability
Significant negative predictors	<ul style="list-style-type: none"> • Practical difficulties with response (identified at first issue stage) • Particular Police Force Areas (bottom 3 = Avon & Somerset, Thames Valley and Metropolitan /City of London) • Low first issue contact probability
Nagelkerke R ²	6.4%

Summary

This last model covers both contact and co-operation rates across multiple re-issue stages. Cases may be re-issued up to four times although only a tiny fraction are (see Table 10.5), given the need to complete fieldwork assignments within a four month timeframe. Another reason why multiple re-issues is rare is the low productivity rate. Only 20% of cases re-issued for the first time prove productive. This rate drops to between 12% and 14% at subsequent stages.

Table 10.5: Number of cases issued at each stage and productivity rates

Original issue	66,387 cases: 17,457 unproductive non-deadwood (29% of issued non-deadwood)
1 st re-issue	11,727 cases (67% of available): 9,349 unproductive non-deadwood (80% of issued)
2 nd re-issue	3,418 cases (37% of available): 2,929 unproductive non-deadwood (86% of issued)
3 rd reissue	730 cases (25% of available): 634 unproductive non-deadwood (87% of issued)
4 th reissue	40 cases (6% of available): 35 unproductive non-deadwood (88% of issued)

Re-issues also stretch fieldwork resources in a geographically uneven fashion which means that unproductive cases may be more or less likely to be re-issued depending on local resource levels. This is best illustrated by looking at the proportion of unproductive original issue cases that were included at the 1st reissue stage. In total this was 67% but it was only 51% in London (the rate in other regions ranged from 65% to 77%).

All of this complicates any model of re-issue response since it will comprise genuine differences in contact and co-operation probabilities (which may occasionally work against each other, obscuring the influence of each) plus differences in resource availability. Attempting to control for all of these factors is virtually impossible and would, in any case, lead to extremely weak models. As it is, this single omnibus model accounts for only 6.4% of the variance in re-issue response levels.

The only location-based variable that proves significant is Police Force Area, and this is probably due to the local variation in re-issue rate rather than anything else.

The two more interesting variables are derived from original issue outcome data.

The first shows the different outcome groups and shows quite clearly that the most productive re-issues are non-contacts. Controlling for the other variables in the model, the probability of achieving an interview from an original household non-contact is twice that of an original refusal. However, the probability is still low (18% compared to 9%).

The second outcome variable uses the response propensities generated by model 1 (household contact, original issue). It is nowhere near as powerful as the variable detailing the original outcome group but does not overlap it as much as might be expected. It also has an interesting distribution: only the lowest decile seems to imply a higher than average probability of failure at the re-issue stage.

Table 10.6 shows the base odds of achieving an interview at the re-issue stage as well as odds ratios associated with each category of each predictor. Both the base odds and the odds ratios are shown in the column labelled 'Exp(B)'. The next two columns show the 95% confidence intervals for these statistics. Large intervals generally reflect small sample sizes while small intervals generally reflect large sample sizes.

The overall odds for a particular combination of categories is computed by taking the base odds and multiplying it by the product of the category odds ratios. The predicted contact probability follows via this equation: $\text{probability} = \text{odds} / (\text{odds} + 1)$.

Table 10.6: Odds ratios for re-issue success

Parameter	Exp(B)	95% Confidence Interval for Exp(B)	
		Lower	Upper
BASE ODDS	0.10	0.07	0.13
ORIGINAL ISSUE OUTCOME GROUP			
Ill/hospital/incapable/no English/other unsuccessful - assume majority PROXY	0.73	0.59	0.90
Household contact but no sampled individual	0.97	0.82	1.15
<i>Refusal by sampled individual</i>	<i>1.00</i>	.	.
Sampled individual but no contact with this individual	1.52	1.27	1.81
Agrees to do interview but does not do it/contact but no specific appointment	1.94	1.63	2.31
No household contact	2.21	1.86	2.63
HOUSEHOLD CONTACT DECILE (OI)			
Lowest	0.67	0.53	0.86
2 nd	0.80	0.62	1.02
3 rd	0.88	0.70	1.11
4 th	0.88	0.70	1.12
<i>5th</i>	<i>1.00</i>	.	.
6 th	1.11	0.85	1.44
7 th	1.00	0.79	1.28
8 th	0.95	0.75	1.22
9 th	1.24	0.97	1.58
Highest	1.00	0.78	1.28
POLICE FORCE AREA			
Avon & Somerset	0.91	0.53	1.56
Thames Valley	0.92	0.53	1.60
<i>Metropolitan/City of London</i>	<i>1.00</i>	.	.
Leicestershire	1.06	0.66	1.69
North Yorkshire	1.12	0.69	1.83
Devon & Cornwall	1.17	0.73	1.86
Lancashire	1.20	0.81	1.78
Cheshire	1.21	0.76	1.94
Warwickshire	1.31	0.87	1.97
Hertfordshire	1.34	0.89	2.02
Sussex	1.39	0.91	2.13
Greater Manchester	1.47	1.02	2.14
South Wales	1.50	0.94	2.42
Humberside	1.55	1.08	2.22
Staffordshire	1.56	1.07	2.28
Dorset	1.57	0.99	2.49
Cambridgeshire	1.64	1.16	2.32
Surrey	1.72	1.03	2.86
Wiltshire	1.72	1.08	2.74

North Wales	1.72	1.19	2.50
West Mercia	1.73	1.22	2.43
Nottinghamshire	1.76	1.21	2.54
South Yorkshire	1.79	1.17	2.74
West Midlands	1.85	1.26	2.69
Bedfordshire	1.86	1.43	2.43
Northamptonshire	1.97	1.30	2.99
Gwent	2.02	1.39	2.93
Hampshire	2.05	1.32	3.20
West Yorkshire	2.07	1.40	3.06
Essex	2.10	1.37	3.22
Lincolnshire	2.12	1.49	3.01
Merseyside	2.21	1.39	3.52
Kent	2.36	1.59	3.52
Cumbria	2.39	1.62	3.54
Cleveland	2.43	1.82	3.25
Norfolk	2.45	1.55	3.88
Suffolk	2.55	1.82	3.56
Derbyshire	2.60	1.74	3.88
Gloucestershire	2.62	1.86	3.70
Northumbria	2.63	1.72	4.02
Dyfed Powys	2.70	1.71	4.26
Durham	4.42	3.06	6.37

WEIGHTING

Basic weighting strategy

These models all produce response propensity scores for each case. Consequently, it is possible to produce a compound response propensity. The inverse of this value may be used as a weight to reflect differential levels of non-response. However, a simpler option is to take the set of variables identified as predictive of at least one of (a) original contact rate, (b) original co-operation rate and (c) reissue productivity rate and use these as predictors in a single model of response at any stage.³⁰

Currently, the BCS employs three weights: a design weight, a two-category non-response weight (the 'inner city' weight) and an additional 'calibration' weight to force certain distributions (sex, age and region) to match known population parameters. Our intention is to replace the inner city weight with a non-response weight generated by this model.

The predictive variables used in the model are:

- Presence/absence of entry phone
- Whether/not in neighbourhood watch area
- Accommodation type
- Housing condition
- OAC
- PFA

This new three stage weight should incorporate more non-response information than the current weight and, consequently, may lead to a reduction in non-response bias.

Some cases are missing data for at least one of the four interviewer observations. For these cases, we recommend *imputing* weights on the basis of the available information. We have used a regression-based imputation method here.

³⁰ An alternative is to use these variables as inputs to the SPSS Answer Tree program³⁰. This would produce weighting classes with maximum variance in response rate. The inverse of the class response rate may be used as the non-response weight. However, in our experience the 'models' produced by Answer Tree algorithms are less stable than the models produced by logistic regression. Sometimes only small differences in sample composition can produce quite different weighting classes. Logistic regression parameters also vary from sample to sample but the analyst has much greater control over the model structure.

Weight trimming

The weights were generated and trimmed at the 99th percentile to reduce variance inflation caused by a small number of very large weights. This is standard survey practice since the introduction of a small degree of bias (via trimming) is usually more than made up for by the reduction in variance.

Practical application

The main practical problem is that the BCS datasets are based on the date of *interview* rather than date of issue and therefore include a mix of cases from complete and incomplete fieldwork assignments. However, data from complete assignments is required to generate non-response weights.

This problem can be overcome by using data from the most recent complete four quarters of assignments to derive weights for each cell in the six-category cross-tabulation. This suggestion rests on the reasonable assumption that the mechanisms driving response do not fluctuate greatly from quarter to quarter.

For example, the non-response weights for the July/August/September 2009 dataset could be derived from an analysis of cases issued between April 2008 and March 2009. The weights may be updated each quarter in this way but without revising already issued datasets.

The precise steps to take are included in the appendix.

The effect of weighting

The weights have been applied to a dataset based on all *issued* cases from the 2008-9 survey year. This is different from the official 2008-9 dataset because it *excludes* cases interviewed during the April 08 to March 09 period but sampled from the previous survey year and *includes* cases interviewed outside of this period but sampled for the 2008-9 survey year.

The first thing to check is how far the new non-response weight is correlated with the 'inner city' weight it is intended to replace. A very low correlation would mean a risk of serious disjuncture in the time series if/when the new weight replaces the old

weight. On the other hand, a very high correlation would cast doubt on the value of replacing the old weight.

The observed correlation was .242, statistically significant but not high. This can be calculated for each quarter to see how much this might vary from quarter to quarter. It varied from .174 to .293, a small enough range to suggest a fairly stable effect. An extension of this is to check how far this correlation varies from year to year. In 2006-07 it would have been .224; in 2007-8 it would have been .306. This is also a small range.

A second thing to check is how stable the parameter estimates are from year to year. This is complicated by the fact that OA group was not appended to previous datasets. ACORN group has had to stand in its stead. On the whole, we observe strong correlations in the parameters between years, albeit with some instability in a small number of parameters due to small sample sizes (e.g. rare accommodation types). This was expected and not a major problem as far as weighting is concerned. It would have been more of a problem if parameter values had fluctuated significantly for large sample categories (they do not).

The third thing to check is how far the two rival weights are correlated with substantive BCS variables. The obvious start-point is the summary variable <victim>, a simple binomial variable tagging respondents as either a victim of crime or not a victim of crime.

Here the new non-response weight outperforms the old inner city weight. The correlation with <victim> is .076 compared to .027. Neither coefficient is large but both are significantly different from zero at the 99% level. The new weight also appears to have a slightly more stable relationship with <victim>. Split by quarter, the correlation between the new weight and <victim> ranges from .069 to .089. The correlation between the old weight and <victim> ranges from .011 to .049.

This makes a small but noticeable difference to the weighted proportion classified as a victim. With the old weight it is 25.8%, with the new weight it is 26.1% (+0.3%). This difference ranges from 0.1 percentage point to 0.4 percentage points when the data is split by quarter.

This evidence suggests that the new weight is better than the old weight at reducing non-response bias although the effect of the final stage – calibration to population totals – might reduce some of this advantage.

However, the other side of the error equation is variance. If the new weights hold greater variance than the old weights then adoption of the new weights may not be advisable, whatever the bias reduction.

The easiest way of comparing the two is to use the shorthand ‘design’ estimate of variance which presumes no design effects due to clustering or stratification. This analysis suggests that the impact of the non-response weight on variance is trivial. The design effect associated with the old inner city weight is 1.001; with the new non-response weight it is 1.007.

Consequently, we conclude that the new weight reduces error. However, the final test of this will follow the calibration weight procedure.

Although a new set of ‘shadow’ weights for the 2009-10 survey have been computed as described above because the testing process was not yet completed it was decided to use the existing weighting schema as previous surveys for the 2009-10 results. Consequently any new weights will not be introduced until 2010-11 at the earliest.

Annex 2 – 2009-10 British Crime Survey: Computation of the under 16s non- response rate

In addition to exploring non-response on the adult survey a non-response weight was calculated for the 10-15 year olds survey.

Analysis of the non-response was carried out using logistic regression to obtain a probability of the child taking part in the survey.

The variables that were found to be significantly associated with non-response were:

- Number of adults in the household from the core adult interview.
- Age of sampled child
- Whether the sampled child had a mobile phone stolen
- Length of core adult interview
- Whether the adult completed the self completion section
- How confident the adult was that the police are effective at catching criminals;
and
- Type of newspaper the adult mainly reads

The following were not significant, but were included for completeness:

- Whether the adult was a victim of crime; and
- Sex of sampled child

Introduction

The BCS child survey aims to sample eligible 10 – 15 year olds in households where an interview with an adult is achieved on the core BCS survey.

In computing the weights, analysis was restricted only to households that were issued as part of the core survey from 1st January 2009 to December 2009.

The first stage was to identify households in the core sample that contained at least one 10 to 15 year old who was eligible for the child survey. These households were considered to be the 'issued sample' for the child survey.

A total of 5,341 households were identified in the January 2009 – December 2009 data file which contained at least one eligible 10 to 15 year old. Of these, an interview was achieved with a 10 to 15 year old in 3,661 households (68.5%).

Using logistic regression the data was analysed to see which variables were associated with non response. A range of variables were initially tested, and those that appeared to be related to non-response were identified and tested further.

The variables that were tested can be grouped into 4 main categories:

- Sample information
- Household information
- Characteristics of the adult respondent; and
- Characteristics of the child

For the purposes of the analysis some questions had to be re-coded. These included the sex and age of the sampled child, as well as allocation of missing values, and banding variables into appropriately sized groups. More details of how the variables were coded are given below.

Design weights that accounted for the probability of selection of an eligible child were applied.

The questions were then reduced to those that were considered to be related to response and tested through CHAID (Chi-Square Automatic Interaction Detector) and using logistic regression.

From these tests the main variables were chosen to form the final model for the non-responding children. A probability of response was then assigned to each child based on their responses to the questions chosen for the model. The probability of response was then inverted to give a weight. This weight was then applied to the responding child, to act as a non-response weight.

The weights from the main adult file were then attached, and multiplied by the child response weights to give an overall weight for non-response at the child and household levels.

Recoding variables

A number of variables had to be re-coded in order for the analysis to be carried out. These included sample variables for the issued child; assigning values to missing cases; and banding categories into suitably sized groups for a number of variables.

In the majority of cases, the number that identified the child that had been sampled was available. In these cases the information from the household grid from the core

survey was used to identify the sex and age of the child, and whether or not the child had had their mobile phone stolen.

There were some cases where the information that indicated which child was sampled was not available. These were all cases where the child interview had not been carried out. In some cases the child screening and selection of an eligible child would have been done by the interviewer and recorded on the Contact Sheet³¹, while in other cases the child screening and selection of an eligible child may not have been carried out.

In cases where there was only one eligible 10 to 15 year old in a household it was simple to work out the child number of the sampled child, and their details were taken from the household grid. Where there was more than one eligible child (around 10% of issued cases), a child number was randomly assigned from the household grid, and the information about that child was then taken from the household grid³².

The remaining data was then recoded into a more manageable form. This mainly involved allocating 'missing' values into the most appropriate groups, and banding categories in to groups of sufficient sample sizes.

³¹ Interviewers were not asked to report this information electronically

³² In essence this is replicating the random selection process that may or may not have been carried out by the interviewer in the field.

Model testing

The table below shows a summary of the main models that were tested using CHAID models and a stepwise logistic regression, using various combinations of the variables listed below. Those variables that consistently came out as significant were included in the final model.

Whether child had mobile phone stolen or not
Interview length (banded)
Sex of interviewer
ONS Harmonised tenure type
Are the contents of your home insured
Respondent de facto marital status
How is your health in general
What newspaper respondent reads
Has respondent accepted self-completion
How confident are you that the police are effective at catching criminals?
How good a job are the police IN THIS AREA doing
The police in this area treat everyone fairly regardless of who they are
The police in this area are dealing with the things that matter to this community
Respondent working in last 7 days
Whether watch news on TV
Month of issue
Month of adult interview
Core survey module completed (A-D)
Number of adults in household (grouped)
Number of children in household (grouped)
Whether adult a victim of crime
ONS harmonised length of time at address
Day of week of adult interview
If something stolen off or out of vehicle
Age of sampled child
Sex of sampled child
Education level of respondents
Respondent NS-SEC
HRP NS-SEC
Problems with noisy neighbours
Problems with rubbish

The final model chosen used logistic regression to obtain the probability of response based on the following variables:

- whether sampled child had mobile phone stolen (no phone; has phone-not stolen; has phone-stolen)
- length of adult interview (banded <1h30, 1h30+)
- Main newspaper readership (broadsheet, Tabloid, other/no main paper, none)
- Whether Adult accepted self completion (Yes, No)
- How confident are you that the police are effective at catching criminals
- Number of adults in the household (1,2,3,4,5+)
- Age of child sampled

The following were not significant, but were included for completeness:

- Whether adult is a victim of crime
- Sex of sampled child.

Appendix 1. shows the response rates for the different groups used for weighting, and their issued eligible samples sizes.

Creating the final weights

There were several steps to creating the final weight for the child. The non-response weight that incorporates the design weight for the number of eligible children in the household (ChildNRwt2d) is based on responding households. The household non-response weight (HHwt) from the core adult file is multiplied by the child non-response weight to give an overall unscaled and untrimmed child weight (UsUtChildJD09wt). This was capped at the 99th percentile so as to reduce the impact of any unusual, large weights (UsTrChildJD09wt99), and then scaled so that the weighted sample size matched that of the achieved sample size (ScTrChildJD09wt99).

Appendix 1.

Questions used in weighting	Response rate	Issued sample size
Overall	68.5%	5341
Sex of child		
Male	68.6%	2760
Female	68.5%	2581
Age of sampled child		
10	66.0%	755
11	70.1%	877
12	70.5%	879
13	68.4%	890
14	71.8%	967
15	64.2%	973
Whether Adult respondents is a victim		
Not victim	68.3%	3679
Victim	69.1%	1662
NS-Sec of Household Reference Person		
Managerial	71.0%	2042
Intermediate	66.5%	1066
Routine/Manual	67.4%	1923
None	66.1%	310
Age of Adult respondent		
16-19	73.0%	488
20-24	72.5%	102
25-34	63.9%	584
35-44	68.3%	2480
45-54	68.7%	1488
55-64	69.5%	151
65-74	73.2%	41
75-84	83.3%	6
85+	100.0%	1

Questions used in weighting	Response rate	Issued sample size
Sex of respondent		
Male	68.4%	2061
Female	68.7%	3280
Whether YP had a mobile phone stolen (according to Household grid)		
Does not own Mobile phone	62.2%	864
Owns Mobile Phone but not stolen	69.6%	4348
Mobile phone stolen	76.7%	129
Length of Adult Interview		
Less than 1h30	68.9%	5048
1h 30 or longer	62.5%	293
Newspaper readership of Adult respondent		
Tabloid	66.6%	2581
Broadsheet	71.7%	923
Other/no particular paper	68.6%	156
None	69.8%	1681
Whether Adult respondent has accepted self completion		
Self completion accepted	69.7%	4388
Self completion refused	60.7%	392
Completed by interviewer	65.1%	561
How confident Adult respondent is that the police are effective at catching criminals		
Very confident	75.0%	376
Fairly confident	69.7%	3046
Not very confident	64.9%	1531
Not at all confident	67.8%	388
Number of Adults in household (grouped)		
One	64.6%	987
Two	68.2%	2902
Three	72.0%	1121
Four	72.4%	279
Five or more	69.2%	52

Questions used in weighting	Response rate	Issued sample size
Number of Children in household (grouped)		
One	68.5%	2214
Two	69.3%	2086
Three	67.7%	756
Four	66.3%	208
Five or more	62.3%	77

