



Economic and Social Data Service

British Crime Survey 2007-2008: Unrestricted Access Teaching Dataset (SN 6891)

ESDS Government

User Guide

Contact:

Sarah King-Hele

ESDS Government

The Cathie Marsh Centre for Census and Survey Research (CCSR)

University of Manchester, Manchester, M13 9PL

Tel: 0161 275 4262; Email: help@esds.ac.uk

Updated: November 2011

Contents

Introduction to the British Crime Survey	3
Historical outlook	3
Aims.....	4
Questionnaire development	5
Sample design	5
Sampling errors and design effects.....	6
Weighting the data	7
The British Crime Survey 2007-2008: Unrestricted Access Teaching Dataset	8
How to obtain the BCS 2007-2008: Unrestricted Access Teaching Dataset	8
The data and variables	8
Weights within the dataset.....	8
Missing values within the dataset.....	9
Variables added to the unrestricted access teaching dataset	9
List of variables	11
Codebook	13
Missing values	13
The dataset.....	13

Introduction to the British Crime Survey

The British Crime Survey (BCS) is one of the largest social surveys conducted in Britain. It 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 the person), which they themselves have experienced in the 12 months prior to the date of interview. The wording of the series of questions which are asked to elicit victimisation experiences have been held constant throughout the series of BCS surveys.

Because members of the public are asked directly about victimisation, the BCS provides a record of the experience of crime which is unaffected by variations in the behaviour of victims about reporting the incident to the police, and variations over time and between places in the police practices about recording crime. 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. The BCS has been successful at developing special measures to estimate the extent of domestic violence, stalking and sexual victimisation, which are probably the least-reported to the police but among the most serious of crimes in their impact on victims.

The BCS data is deposited at the UK Data Archive at the University of Essex, which holds the data for each BCS carried out since 1982. Further information about the methodology of the BCS can be found in the [BCS Technical Report](#) on the UK Data Archive website <www.esds.ac.uk>.

Historical outlook

The BCS has been carried out in England and Wales in 1982, 1984, 1988, 1992, 1994, 1996, 1998, and annually from 2000. Since April 2001, interviewing has been carried out continually and reported on in financial year cycles. It was conducted by a consortium of the National Centre for Social Research (formerly SCPR) and the Office for National Statistics. The 1982 and 1988 surveys were also conducted in Scotland. Users should note that the 1988 Scottish survey was also known as the Scottish Areas Crime Survey. Since 1993, separate Scottish Crime Surveys have been conducted approximately once every three years.

Aims

The main aim of the BCS is to estimate the extent of crime against individuals and their private property. It provides an alternative measure of crime to that provided by the recorded crime statistics. The BCS asks adults aged 16 years and over living in private households in England and Wales about their experiences of criminal victimization over the previous 12 months, regardless of whether or not they reported the incident to the police. To enable comparisons between surveys, the core questions on victimization have remained unchanged since the first sweep.

There are three BCS measures of the extent of crime in England and Wales:

- incident rates: the number of crimes per 10,000 adults (aged 16 plus) or households in England and Wales
- prevalence rates (also known as risks): the percentage chance of an adult or household being a victim once or more
- crime count: the total number of crimes (calculated by applying the incident rates to population figures)

All three measures are calculated for the financial year. So, the 2007-2008 BCS measures crime from April 2007 to March 2008. It should be noted that the figures derived from the BCS are *estimates*. As with any sample survey the BCS estimates are subject to sampling error and a range of other methodological limitations.

The survey includes demographic and lifestyle variables, both for the respondent and the head of household, which allow the identification of relative risks of victimisation by socio-economic, demographic and lifestyle factors. The BCS also collects information on the nature of crime, including where and when offences occur and the emotional, physical and financial impact of crime. On an ad-hoc basis it also covers various other crime-related issues such as concern about crime, attitudes to the police and drug misuse. The survey is also used to measure non-crime issues, such as experiences of fires, on an ad-hoc basis.

For some topics, trend analysis may prove difficult due to the fact that topics are not always covered by the same questions each time. For non-crime sections there is the concern that responses will be affected by the overall crime context of the survey. However, great care is

taken to minimize contextual effects by choosing the most appropriate place in the survey to place non-crime topics.

Questionnaire development

The 2007-2008 BCS interview comprised 6 main sections. These were as follows:

- introductory questions about the household
- screener - questions used to identify victimization experiences
- victim forms for any incidents identified at the screeners (up to a limit of six forms)
- follow-up modules: experience of the police, attitudes to the criminal justice system, crime prevention and disorder, and an ad hoc module
- modules on night-time economy, anti-social behaviour, crime and disorder in town centres and on public transport and domestic violence and sexual victimization
- self-completion sections on drugs and drinking behaviour

The entire interview was administered using computer-assisted personal interviewing (CAPI). All respondents were asked the introductory questions and the screeners, although within these sections particular questions were asked of sub-groups and in some cases two variants of the same questions were each asked of half the sample. Victim Forms were automatically included in the interview by the CAPI program, to collect details of incidents identified at the screeners. A maximum of six Victim Forms could be asked. A long version of the Victim Form was asked for the first three incidents and a shorter version for the fourth to sixth incident and for any that had occurred outside of England and Wales.

The victimization screening questions were designed to ensure that all incidents of crime that fit within the scope of the BCS, including relatively minor ones, are mentioned. The screener questions also aim to ensure that each incident is only counted once.

Sample design

The BCS sample is designed to give, after appropriate weighting, both a representative cross-section of private households in England and Wales, and of individuals aged 16 and over living in them. Since 1992 the Small Users Postcode Address File (PAF) has been used as the sampling

frame. The PAF, listing all postal delivery points in the country, represents the fullest register of household addresses as almost all households have one delivery point, or letterbox.

A stratified multi-stage random probability design is used to select the sample of addresses. As with all large-scale surveys the BCS sample is clustered to keep costs at an acceptable level without compromising the quality of the sample. Since 1992 the procedure has been as follows. Postcode sectors are sorted into 10 standard regions. Within each region, sectors are listed in order of population density and divided into three roughly equal-sized bands, in terms of the number of delivery points. Within each of the 30 strata, sectors are ordered in increasing order of the percentage of households with head of household's socio-economic group (SEG) defined as professional and managerial, other non-manual or skilled manual. By sampling systematically down the ordered list, using a random starting point, postcode sectors are selected with a probability proportional to size (number of PAF addresses). Within each of the postcode sectors selected, the list of delivery points is divided into four equal-sized segments. One of these segments is chosen at random.

Where one address has more than one household, a single household is selected using random selection procedures. One adult aged 16 or over in each selected household is identified for interview using similar random-selection procedures. No substitution of respondents is allowed.

Inner city areas are oversampled by a factor of about two. Inner city areas are selected on the basis of classifying postcode sectors according to population density, level of owner-occupied tenure, and social class profile.

Sampling errors and design effects

If the BCS sample was a simple random sample of dwelling units in Britain, the estimates produced at this stage would represent victimization estimates covering England and Wales for the time period of the survey. However, the sample is clustered within postcode sectors and different individuals are selected with different levels of probability. While weighting removes these differences in selection, the weighted results are not based on a simple random sample. An estimation procedure is used to calculate the extent to which the estimates need to be set within wider confidence intervals, due to the complex nature of the sample design.

Weighting the data

The BCS contains a number of weights that should be applied for different types of analyses. However, for the purpose of the unrestricted access teaching dataset, only the weight for individual-level analyses is available: *tcindwt*.

The British Crime Survey 2007-2008: Unrestricted Access Teaching Dataset (SN 6891)

How to obtain the BCS 2007-2008: Unrestricted Access Teaching Dataset (SN 6891)

The BCS teaching dataset can be obtained from the UK Data Archive and unlike most data available through ESDS *is NOT subject to registration with ESDS*. The data and documentation may also be downloaded from the ESDS Government web-pages:

<https://www.esds.ac.uk/government/resources/teachingdatasets.asp>

The teaching dataset is available in two formats: SPSS and STATA.

Note that there is a version of this teaching data available from the UK Data Archive under the standard End User Licence (EUL). The British Crime Survey 2007-2008: Teaching Dataset (SN 6561) contains a wider range of variables than the related unrestricted access data and has its own user guide. The link to this teaching dataset and documentation can be found here:

<https://www.esds.ac.uk/findingData/snDescription.asp?sn=6561>

The data and variables

The data file contains 11,676 adults aged 16 years and over in England and Wales. This represents roughly a quarter of the core sample in 2007-2008, only those individuals who were randomly assigned to answer the Module B questions. The unrestricted access teaching dataset contains 35 variables.

Weights within the dataset

Different units of analysis (households and individuals) in the BCS dataset have different probabilities of inclusion in the sample. These differences arise from a number of sources: the over-sampling of small Police Force Areas (PFAs), the sub-selection of one dwelling unit at an address, the selection of one individual within a dwelling and differential response rates within subgroups. It is necessary to correct for these differences, by weighting, in order that estimates will be unbiased. Otherwise, the sample would over-represent small PFAs, single-dwelling addresses and people living alone.

The sample is designed to be representative of the entire household population of England and Wales, so use of *tcindwt* will provide nationally-representative estimates. This weight variable is the individual weight provided in the full 2007-2008 BCS (*indivwgt*) divided by its mean value.

All of the variables included within the unrestricted access teaching dataset are individual-level variables, and require individual-based analysis. In interpreting and presenting the analysis the weighted percentages, means etc should be used. Unweighted information is only used for the sample size.

Missing values within the dataset

Respondents on the BCS are not usually explicitly given the options of “refusal” or “don’t know” when asked a question. However, respondents may say they do not know or they may refuse to answer a question or they may feel that the question does not apply to them.

SPSS version: The SPSS teaching dataset has all ‘refusal’ and ‘don’t know’ responses set as missing values. Some variables within the dataset also contain ‘system missing’ cases - SPSS automatically codes an empty cell as ‘system missing’ which is denoted by a dot (.). The codes for missing data for the SPSS version of the data are 7 for ‘not applicable’, 8 for ‘refused’ and 9 for ‘don’t know’ responses.

Stata version: The STATA teaching dataset includes all ‘refusal’ and ‘don’t know’ responses as missing responses. The equivalent of ‘system missing’ data are denoted by a dot (.), while .a denotes ‘not applicable’, .b denotes ‘refused’ and .c denotes ‘don’t know’ responses.

Variables added to the unrestricted access teaching dataset

Note that the unrestricted access teaching dataset contains some variables which were derived from other variables on the BCS. You can identify these variables as they are prefixed ‘tc’. These variables have been added for teaching purposes only and for students to use in analyses and reports as part of a course. All analyses and research for other purposes should be conducted using the original BCS 2007-2008 (SN 6066) dataset available via the ESDS website: <https://www.esds.ac.uk/findingData/snDescription.asp?sn=6066>.

Further notes about the methods used to create the additional 'tc' variables can be found in the document *Notes for Teachers* attached to the British Crime Survey 2007-2008: Teaching Dataset (SN 6561) on the ESDS website. Links to the documentation are at the bottom of the webpage.

<https://www.esds.ac.uk/findingData/snDescription.asp?sn=6561>.

List of variables

The list below contains the 35 variables in the dataset in the order in which they appear in the SPSS or STATA files. Note that the order in which these variables appear has been adjusted from that in the original BCS 2007-2008 to facilitate grouping by topic.

No.	Name	Label	Measure
Case id and household variables			
1	rowlabel	Case identifier (8 digits)	Scale
Socio-demographic variables			
2	sex	Respondent sex	Nominal
3	age	Respondent age	Scale
4	livharm1	Marital status (ONS harmonised)	Nominal
5	ethgrp2	Respondent ethnic origin (5 categories)	Nominal
6	educat3	Respondent education (5 categories)	Nominal
7	work	Any paid work in last week	Nominal
Accommodation and area characteristics			
8	ysarea	How long have you lived in this area?	Scale
9	resyrago	Were you living at this address 12 months ago?	Nominal
10	tenure1	In which way do you occupy this accommodation?	Nominal
11	rural2	Type of area 2004: urban/rural	Nominal
12	rubbcomm	In the immediate area how common is litter\rubbish?	Ordinal
13	vandcomm	How common is vandalism graffiti or damage to property?	Ordinal
14	poorhou	How common are homes in poor condition\run down?	Ordinal
15	tcemdiqu2	Index of multiple deprivation by quintile in England (1=20% most deprived wards)	Scale
16	tcwmdiqu2	Index of multiple deprivation by quintile in Wales (1=20% most deprived wards)	Scale
Fear of crime			
17	causem	One MAIN cause of crime in Britain today	Nominal
18	walkdark	How safe do you feel walking alone after dark?	Ordinal
19	walkday	How safe do you feel walking alone in this area during the day?	Ordinal
20	homealon	How safe do you feel when alone in home at night?	Ordinal
21	tcviolent	Respondent level of worry about being a victim of personal crime (high score=high level of worry)	Scale
22	tcsteal	Respondent level of worry about being a victim of property crime (high score=high level of worry)	Scale
23	wburgl	How worried about having your home broken into?	Ordinal
24	wmugged	How worried about being mugged and robbed?	Ordinal
25	wcarstol	How worried about having car stolen?	Ordinal
26	wfromcar	How worried about having things stolen from your car?	Ordinal

27	wraped	How worried about being raped?	Ordinal
28	wattack	How worried about being physically attacked by strangers?	Ordinal
29	winsult	How worried about being insulted or pestered by anybody?	Ordinal
30	wraceatt	How worried about being attacked because of skin colour?	Ordinal

Respondent opinion about anti-social behaviour and crime in their area

31	crimerat	How much crime rate has changed in this area since 2 years ago?	Ordinal
32	tcarea	Respondent opinion about the level of problems with anti-social behaviour in their neighbourhood (high score=high levels of anti-social behaviour)	Scale
33	tcneigh	Respondent opinion about level of problems with noisy or nuisance neighbours in their neighbourhood (high score=high level of problems with neighbours)	Scale

Experience of crime in the last 12 months

34	bcsvictim	Experience of any crime in the last 12 months	Nominal
----	-----------	---	---------

Weight

35	tcindwt	Weight to be used when analysing individual-level data (mean=1)"	Scale
----	---------	--	-------

Codebook

Missing values

The following shows the frequencies including missing values for categorical variables, and missing values and summary statistics for scalar variables. Missing values appear in different formats in SPSS and STATA; they are summarized in the table below:

Missing value	SPSS	STATA
System missing	.	.
Not applicable	7	.a
refused	8	.b
Don't know	9	.c

The Dataset

1. rowlabel case identifier (8 digits)

type: numeric (long)
range: [61302140,86052180] units: 10
unique values: 11676 missing .: 0/11676
mean: 7.6e+07
std. dev: 4.2e+06
percentiles: 10% 25% 50% 75% 90%
7.3e+07 7.3e+07 7.5e+07 7.6e+07 8.4e+07

2. sex respondent sex

type: numeric (byte)
label: LABB
range: [1,2] units: 1
unique values: 2 missing .: 0/11676
tabulation: Freq. Numeric Label
5307 1 male
6369 2 female

3. age respondent age

type: numeric (int)
range: [16,101] units: 1
unique values: 84 missing .: 0/11676
unique mv codes: 2 missing .*: 15/11676
mean: 50.4228
std. dev: 18.5389
percentiles: 10% 25% 50% 75% 90%
26 36 49 65 76

 4. livharm1 ons harmonised marital status

type: numeric (byte)
 label: LIVHARM

range: [1,6] units: 1
 unique values: 6 missing .: 8/11676

tabulation:	Freq.	Numeric	Label
	5559	1	married
	1032	2	cohabiting
	2374	3	single
	323	4	separated
	1056	5	divorced
	1324	6	widowed
	8	.	.

 5. ethgrp2 ethnic group (5 categories)

type: numeric (byte)
 label: ethgrp2

range: [1,5] units: 1
 unique values: 5 missing .: 3/11676

tabulation:	Freq.	Numeric	Label
	10900	1	white
	56	2	mixed
	380	3	asian or asian british
	203	4	black or black british
	134	5	chinese or other
	3	.	.

 6. educat3 respondent education (5 categories)

type: numeric (byte)
 label: educat3

range: [1,5] units: 1
 unique values: 5 missing .: 58/11676

tabulation:	Freq.	Numeric	Label
	3320	1	none
	2303	2	o level/gcse
	1987	3	apprenticeship or a/as level
	3518	4	degree or diploma
	490	5	other
	58	.	.

 7. work any paid work in last week

type: numeric (byte)
 label: LABM

range: [1,2] units: 1
 unique values: 2 missing .: 33/11676
 unique mv codes: 3 missing .*: 3/11676

tabulation:	Freq.	Numeric	Label
	6345	1	yes
	5295	2	no
	33	.	.
	2	.b	.b
	1	.c	.c

8. yrsarea how long have you lived in this area

type: numeric (byte)
label: LABK
range: [1,7] units: 1
unique values: 7 missing .: 0/11676

tabulation:	Freq.	Numeric	Label
	740	1	less than 12 months
	600	2	12 months but less than 2 years
	577	3	2 years but less than 3 years
	893	4	3 years but less than 5 years
	1660	5	5 years but less than 10 years
	2061	6	10 years but less than 20 years
	5145	7	20 years or longer

9. resyrago were you living at this address 12 months ago

type: numeric (byte)
label: LABM
range: [1,2] units: 1
unique values: 2 missing .: 9662/11676

tabulation:	Freq.	Numeric	Label
	814	1	yes
	1200	2	no
	9662	.	

10. tenure1 in which way do you occupy this accommodation

type: numeric (byte)
label: tenure1
range: [1,6] units: 1
unique values: 6 missing .: 33/11676
unique mv codes: 3 missing .*: 37/11676

tabulation:	Freq.	Numeric	Label
	4048	1	own it outright
	4084	2	buying it with the help of a mortgage or loan
	70	3	pay part rent and part mortgage (shared ownership)
	3078	4	rent it
	326	5	live here rent free (inc. rent free in relative/friend's)/squatting
	33	.	
	24	.b	
	13	.c	

11. rural2 type of area 2004: urban/rural

type: numeric (byte)
label: rural2
range: [1,2] units: 1
unique values: 2 missing .: 0/11676

tabulation:	Freq.	Numeric	Label
	8702	1	urban
	2974	2	rural

12. rubbcomm in the immediate area how common is litter\rubbish

type: numeric (byte)
label: LABBS
range: [1,4] units: 1
unique values: 4 missing .: 611/11676

tabulation:	Freq.	Numeric	Label
	204	1	very common
	1244	2	fairly common
	4154	3	not very common
	5463	4	not at all common
	611	.	

13. vandcomm how common is vandalism graffiti or damage to property

type: numeric (byte)
label: LABBS
range: [1,4] units: 1
unique values: 4 missing .: 597/11676

tabulation:	Freq.	Numeric	Label
	79	1	very common
	417	2	fairly common
	3352	3	not very common
	7231	4	not at all common
	597	.	

14. poorhou how common are homes in poor condition\run down

type: numeric (byte)
label: LABBS
range: [1,4] units: 1
unique values: 4 missing .: 628/11676

tabulation:	Freq.	Numeric	Label
	81	1	very common
	604	2	fairly common
	4046	3	not very common
	6317	4	not at all common
	628	.	

15. tcemdiq2 Index of multiple deprivation by quintile in
England (1=20% most deprived wards)

type: numeric (float)
range: [1,5] units: 1
unique values: 5 missing .: 1046/11676

tabulation:	Freq.	Value
	1843	1
	1887	2
	2226	3
	2338	4
	2336	5
	1046	.

 16. tcwmdiqu2 Index of multiple deprivation by quintile in
 Wales (1=20% most deprived wards)

type: numeric (float)
 range: [1,5] units: 1
 unique values: 5 missing .: 10630/11676

tabulation: Freq. Value
 153 1
 228 2
 194 3
 251 4
 220 5
 10630 .

 17. causem one main cause of crime in Britain today

type: numeric (byte)
 label: causem
 range: [1,10] units: 1
 unique values: 10 missing .: 0/11676
 unique mv codes: 1 missing .*: 140/11676

examples: 4 d. lack of discipline from parents
 4 d. lack of discipline from parents
 5 e. drugs
 6 f. alcohol

 18. walkdark how safe do you feel walking alone after dark

type: numeric (byte)
 label: LABP
 range: [1,4] units: 1
 unique values: 4 missing .: 0/11676
 unique mv codes: 1 missing .*: 51/11676

tabulation: Freq. Numeric Label
 3002 1 very safe
 4718 2 fairly safe
 2604 3 a bit unsafe
 1301 4 very unsafe
 51 .c

 19. walkday how safe do you feel walking alone in this area during the day

type: numeric (byte)
 label: walkday
 range: [1,4] units: 1
 unique values: 4 missing .: 0/11676
 unique mv codes: 2 missing .*: 12/11676

tabulation: Freq. Numeric Label
 8488 1 very safe
 2745 2 fairly safe
 369 3 a bit unsafe
 62 4 or very unsafe
 2 .b
 10 .c

24. wmugged how worried about being mugged and robbed

type: numeric (byte)
label: LABR

range: [1,4] units: 1
unique values: 4 missing .: 0/11676
unique mv codes: 2 missing .*: 19/11676

tabulation:	Freq.	Numeric	Label
	1278	1	very worried
	2864	2	fairly worried
	5451	3	not very worried
	2064	4	not at all worried
	14	.a	
	5	.c	

25. wcarstol how worried about having car stolen

type: numeric (byte)
label: LABR

range: [1,4] units: 1
unique values: 4 missing .: 0/11676
unique mv codes: 2 missing .*: 2515/11676

tabulation:	Freq.	Numeric	Label
	979	1	very worried
	2611	2	fairly worried
	3978	3	not very worried
	1593	4	not at all worried
	2514	.a	
	1	.c	

26. wfromcar how worried about having things stolen from your car

type: numeric (byte)
label: LABR

range: [1,4] units: 1
unique values: 4 missing .: 2514/11676
unique mv codes: 2 missing .*: 18/11676

tabulation:	Freq.	Numeric	Label
	828	1	very worried
	2653	2	fairly worried
	4003	3	not very worried
	1660	4	not at all worried
	2514	.	
	18	.a	

27. wraped how worried about being raped

type: numeric (byte)
label: LABR

range: [1,4] units: 1
unique values: 4 missing .: 0/11676
unique mv codes: 3 missing .*: 488/11676

tabulation:	Freq.	Numeric	Label
	1141	1	very worried
	1220	2	fairly worried
	3669	3	not very worried
	5158	4	not at all worried
	471	.a	
	1	.b	
	16	.c	


```

-----
32. tcarea      Respondent opinion about the level of anti-social behaviour IN THEIR
                NEIGHBOURHOOD (high score=high levels of anti-social behaviour)
-----

        type: numeric (float)
        range: [-2.6734681,4.1882639]          units: 1.000e-11
unique values: 7358                          missing  .: 677/11676

        mean: .030254
        std. dev: 1.01006

        percentiles:      10%      25%      50%      75%      90%
                          -1.15729  -.794439  -.094174  .642037  1.40043
-----

33. tcneigh     Respondent opinion about the level of problems with noisy or nuisance
                neighbours IN THEIR NEIGHBOURHOOD (high score=high level of problems with
                neighbours)
-----

        type: numeric (float)
        range: [-2.0767403,4.6414561]          units: 1.000e-11
unique values: 7358                          missing  .: 677/11676

        mean: -.012718
        std. dev: .998659

        percentiles:      10%      25%      50%      75%      90%
                          -.834139  -.600858  -.430028  .324986  1.18496
-----

34. bcsvictim   experience of any crime in the previous 12 months
-----

        type: numeric (byte)
        label: bcsvictim

        range: [0,1]                          units: 1
unique values: 2                              missing  .: 0/11676

        tabulation:  Freq.  Numeric  Label
                    9318    0      not a victim of crime
                    2358    1      victim of crime
-----

35. tcindwt     Weight to be used when analysing individual-level data (mean=1)
-----

        type: numeric (float)
        range: [.17144209,6.1711383]          units: 1.000e-08
unique values: 11221                          missing  .: 16/11676

        mean: 1.00102
        std. dev: .711173

        percentiles:      10%      25%      50%      75%      90%
                          .345566  .50963  .792767  1.33736  1.86161

```