Proposed changes to BCS sample design

BACKGROUND

The current sample design for the British Crime Survey (BCS) yields interviews with a nationally representative sample of 46,000 households in England and Wales each year. With the exception of the City of London police force area (which for the purpose of analysis is merged with the Metropolitan Police Force Area), the sample is designed to yield a minimum of 1,000 interviews with adults (aged 16 years and over) in each one of the remaining 42 territorial police force areas (PFAs).

The requirement for a minimum sample of 1,000 interviews in each PFA was introduced in 2004 and was driven by the introduction of the Police Performance and Assessment Framework (PPAF), which used the BCS to monitor specific performance targets set by the Government. It should be noted that this is a minimum required sample and the actual sample in the more populous PFAs is higher (e.g. 3,984 adults were sampled in the London area and over 1,370 in Greater Manchester and the West Midlands PFAs in the 2010/11 BCS).

Both the abolition of central targets and the outcome of the Government's 2010 Comprehensive Spending Review (CSR) have led to a review of the utility of the BCS sample. The 2010 CSR resulted in a settlement that required a cut of 23 per cent in real terms in Home Office programme and administration costs over the life of the spending period (2012/12 to 2014/15).

The main driver of BCS costs are interviewer fees and expenses, which make up over 90% of the total budget. Thus to achieve cuts of the order required would need a reduction in the actual sample size, a switch to cheaper modes of interviewing (e.g. telephone, internet), or a combination of the two.

FEASIBILITY OF USING CHEAPER DATA COLLECTION MODES

Earlier this year, the Home Office commissioned the current BCS contractor (TNS-BMRB) to examine the feasibility of moving the BCS towards a mixed-mode panel design in which a substantial number of adults were interviewed initially face-to-face (as now) with a sub-sample re-contacted subsequently by a cheaper mode (e.g. via mail, phone or internet) to top-up the sample. Such an approach is used on surveys such as the Labour Force Survey where an initial face-to-face interview (wave 1) is followed up with four subsequent re-contact interviews by telephone at quarterly intervals.

TNS-BMRB designed a series of experiments in which previous BCS respondents were re-contacted by various modes. This allowed the examination of the likely attrition rates and possible non-response bias that might arise from such a design and the possible mode effects from asking questions using different methods. A full report of this work has been published alongside this consultation:

http://www.homeoffice.gov.uk/science-research/research-statistics/crime/crime-statistics/bcs-methodology/

In summary, the work demonstrated that the additional costs of re-contacting respondents could not be justified as the additional error associated with mixing modes would likely cancel out any improvements in precision arising from the increased sample size. Specifically problems were associated with:

- the attrition from initial face-to-face sample through the various stages of re-contact (e.g. permission to re-contact, permission to obtain contact details, contact and co-operation) leading to response rates around half that of the face-to-face sample;

- evidence that the profile of those responding online was systematically different from the face-to-face and other samples; and,

- evidence of mode effects particularly on attitudinal questions (e.g. confidence in the police) across many of the modes with those conducted via self-completion (post and internet) more positive than those by an interviewer.

In addition, the Home Office recognised that such an approach would present logistical difficulties with regard to passing contact details between different contractors without specific permission to do so – the asking of which, is likely to increase non-response further.

It was therefore concluded that it would not be feasible to pursue this option without significant risk to data quality and to continuity trends in BCS crime.

PROPOSED WAY FORWARD

Given the above, it is proposed that the BCS sample size is reduced, from 2012/13 onwards, to an annual achieved sample of 35,000 households per year with one adult randomly selected for a face-to-face interview as now. It is proposed to retain the same interview length as now with retention of the existing self-completion modules on drug and intimate violence. It is also proposed to continue to sample one child aged 10-15 years within households participating in the main BCS. This is expected to yield a nationally representative sample of around 3,100 children in that age range annually (a reduction from around 3,700 currently obtained).

The proposed design should not introduce any discontinuity to key BCS estimates or trends but there will be some loss of precision on those estimates and on the ability to detect statistically significant changes. Under the proposed design, the target is to for a minimum sample of 650 adults per year interviewed in each PFA. While this will inevitably lead to some reduction in the precision of estimates at PFA level this will be fairly modest and we believe it will preserve the ability to produce estimates with a reasonable level of precision at such a level. An example of the confidence intervals under the current and proposed design is given at Annex A for each PFA.

This shows that there would be a fairly small increase in the confidence interval around a typical estimate of public perceptions of the police at a national level. For example, the 2010/11 BCS yields an estimate for the proportion of people in England and Wales who agree that the police and local council are dealing with the crime and ASB issues that matter locally of 52% with a confidence interval of +/- 0.6%. This means that we can be confident (at the 95% level) that the real population value lies between 51.4% and 52.6%. Under the proposed design the confidence interval would increase and the true population value range between 51.3% and 52.7%. Inevitably there will be a larger effect for analysis of particular sub-groups (e.g. victims of specific crime types or users of specific drugs) where base sizes would be smaller.

At regional level (and for Wales as a whole) the effect of the change would be similarly small with the margin of error increasing around this example estimate between 0.1 and 0.5 percentage points. At PFA level the effect of the change will result in increases of the margin of error around survey estimates of around one percentage point or less in most forces. In general, the effect will be smaller on forces covering largely urban areas (e.g. Greater Manchester and West Midlands) and larger on those covering more rural, or less populous, areas (e.g. Bedfordshire, North Yorkshire and Surrey).

To mitigate the effect of a reduction in the actual sample size, a number of innovations are planned which will improve the efficiency of the sample design and these are described below.

Employment of a new address sample design

The proposed design employs a different approach to address sampling. In particular:

- a new bespoke sampling geography for the BCS that minimises the variation in victimisation rates between primary sampling units and thus maximises cluster sample precision; and,

- an evolution of the area sample design in which primary sampling units (PSUs) will be divided into equal thirds on the basis of an improved address density measure. In the 'high density' third, all PSUs will be covered over a one-year period (i.e. the annual sample will be unclustered); in the 'mid-density' third, all PSUs will be covered over a two-year period; and in the 'low density' third, all PSUs will be covered over a three-year period.

The new bespoke sampling geography for the BCS will increase the precision of cluster sample estimates. Minimising variation in primary sampling unit victimisation rates is a crucial requirement if a

cluster sample is to be statistically efficient. The problem is that super output areas which are currently used as PSUs (and, to a lesser extent, postcode sectors) have a neighbourhood-like quality which works against this objective.

It is proposed that this geography will be created by modelling victimisation rates at a small area level (lower level super output areas or LSOAs)) and constructing PSUs from these building blocks, in such a way that the aggregated victimisation rates are as close as possible to the average for the PFA. It has been previously observed that victimisation rates vary most between neighbourhoods when the population density is high and vary least when the population density is low. Consequently, it is the intention to combine together a large number of LSOAs in densely populated areas, while only combining together a small number in sparsely populated areas. The net effect is to standardise any remaining cluster effects so that they do not disproportionately impact upon the more densely populated parts of the country.

LSOAs were chosen as the building block for these new primary sampling units because their small size gives significant flexibility in the construction of these primary sampling units. In addition, there are numerous statistics available about LSOAs which will help with the estimation of victimisation rates. Given known correlations between neighbourhood data and victimisation rates, it is expected that the following variables will be included in the prediction model:

- Output Area Classification profile (a segmentation based on the 2001 census)

- Deprivation index of crime and disorder (updated periodically from CLG/Welsh Government)

- Population density (estimated from address density)

- Tenure profile (from the 2001 census)

- Ethnic mix (from the 2001 census).

Evolution of the area sample design

In addition to the creation of a bespoke BCS sampling geography, a subtle revision of the sampling process is proposed so that the principles that underpin the current design are more smoothly applied. In particular, it is proposed to enforce greater structure on the sample so that, after three years, it will be possible to create a three-year rolling national dataset in which the entire sample is unclustered. Given the reduction in sample size, especially within PFAs, the value of multiple year samples is increased and greater use of them is anticipated.

Another feature of this design is that the London sample will be entirely unclustered each year (compared to 75-80% as in the current design). In a small number of other highly urbanised PFAs (Greater Manchester, Merseyside, the West Midlands and Cleveland) the sample will be unclustered over any consecutive two-year period.

QUESTIONS FOR CONSIDERATION

1)   What are your views on the proposed changes to the BCS sample design?

2)   What impact, if any, will this change have on your use of the BCS?

3)   Do you have any alternative proposals to cut costs?

Responses to this consultation should be sent to the address below (by post or email) by 12 January 2012. Individual responses may be published unless respondents request anonymity.

ANNEX A: EFFECT OF REDUCTION IN SAMPLE ON PRECISION OF ESTIMATES

Estimated impact of sample size changes on estimates of attitudes to local police by police force area, English region and Wales

Percentages   **England and Wales, 2010/11 BCS**

Police force area, English region and Wales	Police and local council are dealing with issues[1]			Impact of proposed new sample design		
	2010/11 survey unweighted base (Number)	2010/11 survey estimate: proportion saying strongly agree/ tend to agree (%)	Margin of error around estimate (+/-)[2]	*Proposed new unweighted base (Number)*	Proposed new margin of error around estimate (+/-)[2]	Difference in margin of error around estimate[3]
Cleveland	1,003	59	3.4	650	4.2	0.8
Durham	994	53	3.4	650	4.2	0.8
Northumbria	1,046	57	3.5	792	4.0	0.5
North East Region	**3,043**	**57**	**2.3**	**2,092**	**2.8**	**0.5**
Cheshire	1,042	54	3.7	650	4.6	1.0
Cumbria	978	50	4.2	650	5.1	0.9
Greater Manchester	1,373	51	2.8	1,434	2.7	-0.1
Lancashire	1,039	55	3.7	786	4.2	0.6
Merseyside	932	55	4.1	750	4.5	0.5
North West Region	**5,364**	**53**	**1.6**	**4,270**	**1.8**	**0.2**
Humberside	1,023	48	4.3	650	5.4	1.1
North Yorkshire	1,002	56	4.5	650	5.6	1.1
South Yorkshire	938	47	4.3	713	5.0	0.6
West Yorkshire	1,099	52	4.0	1,173	3.8	-0.1
Yorkshire and the Humber Region	**4,062**	**50**	**2.2**	**3,186**	**2.5**	**0.3**
Derbyshire	1,002	54	4.4	650	5.4	1.1
Leicestershire	973	55	3.9	650	4.7	0.9
Lincolnshire	993	49	3.8	650	4.7	0.9
Northamptonshire	954	51	3.8	650	4.6	0.8
Nottinghamshire	1,069	49	3.5	650	4.5	1.0
East Midlands Region	**4,991**	**52**	**1.8**	**3,250**	**2.2**	**0.4**
Staffordshire	942	53	4.5	650	5.4	0.9
Warwickshire	1,034	47	2.9	650	3.6	0.8
West Mercia	953	50	4.1	650	5.0	0.9
West Midlands	1,374	50	3.8	1,373	3.8	0.0
West Midlands Region	**4,303**	**50**	**2.2**	**3,323**	**2.5**	**0.3**
Bedfordshire	951	45	5.2	650	6.3	1.1
Cambridgeshire	983	42	4.2	650	5.1	1.0
Essex	991	53	3.5	908	3.6	0.2
Hertfordshire	1,028	53	3.3	650	4.2	0.9
Norfolk	957	52	4.2	650	5.1	0.9
Suffolk	910	53	3.6	650	4.3	0.7
East of England Region	**5,820**	**51**	**1.7**	**4,158**	**2.0**	**0.3**
Metropolitan/City of London	3,984	56	1.5	3,861	1.6	0.0
London Region	**3,984**	**56**	**1.5**	**3,861**	**1.6**	**0.0**
Hampshire	1,009	48	4.2	989	4.2	0.0
Kent	990	51	4.4	888	4.6	0.2
Surrey	1,009	58	4.4	650	5.4	1.1
Sussex	972	50	4.5	856	4.8	0.3
Thames Valley	1,153	54	4.7	1,129	4.7	0.0
South East Region	**5,133**	**52**	**2.0**	**4,512**	**2.2**	**0.1**
Avon & Somerset	951	53	3.6	836	3.8	0.2
Devon & Cornwall	1,014	56	3.3	929	3.4	0.1
Dorset	975	54	3.7	650	4.5	0.8
Gloucestershire	956	53	4.6	650	5.6	1.0
Wiltshire	973	57	4.6	650	5.6	1.0
South West Region	**4,869**	**55**	**1.7**	**3,715**	**2.0**	**0.3**
England Total	**41,569**	**53**	**0.6**	**32,367**	**0.7**	**0.1**
Dyfed Powys	853	52	4.5	650	5.1	0.6
Gwent	1,041	46	3.6	650	4.5	0.9
North Wales	976	46	3.8	650	4.7	0.9
South Wales	954	45	5.1	684	6.1	0.9
Wales	**3,824**	**46**	**2.5**	**2,634**	**3.0**	**0.5**
ENGLAND AND WALES	**45,393**	**52**	**0.6**	**35,001**	**0.7**	**0.1**

1. Respondents are asked how much they agree/disagree with the statement 'The police and local council are dealing with the anti-social behaviour and crime issues that matter in this area'.
2. The margin of error is shown at the 95% level. Adding and subtracting the margin of error from an estimate gives the confidence interval.
3. The difference in confidence intervals shows how much more uncertain survey estimates will be owing to changes in sample sizes.