



BARB 2010 Service

BARB Reference Manual

1. Establishment Survey

VERSION 2.2

Table of Contents

1. ESTABLISHMENT SURVEY	
1.1 Introduction	3
1.2 Establishment Survey Design	3
1.3 Establishment Survey Fieldwork	8
1.4 Establishment Survey Editing Procedures	8
1.5 Establishment Survey Weighting	9
Appendix – Sampling Area Segment Definition	13

1. ESTABLISHMENT SURVEY

1.1 Introduction

Establishment Surveys are conducted on a continuous basis throughout the year. The Establishment Survey has two main objectives:

- (i) To monitor the characteristics of UK households with respect to personal and household demographics, and the ownership and use of television viewing equipment. The data are used to project universe characteristics for the purposes of panel control and the reporting of audiences.
- (ii) To provide a sampling frame of households (the potentials file) from which the meter panels are constructed.

The Establishment Survey is selected as an annual stratified random probability sample of households, which is then divided into twelve monthly replicates. The design is such that the monthly samples are representative of the Network reporting and the quarterly samples representative of ITV Areas and BBC Regions. In-home face-to-face interviews are conducted using CAPI and the results are grossed up to Household and Individuals population estimates derived from Government estimates.

1.2 Establishment Survey Sample Design

1.2.1 Overview

The Establishment Survey Sample will be a stratified probability sample of private households in the UK. The interviewer assignment consists of addresses selected from the Postcode Address File, producing an average of 12 interviews per assignment. The sample operates as 12 monthly waves per year. Samples need to balance at different levels for monthly, quarterly and annual reporting.

RSMB are responsible for the specification and review of the core Establishment Survey Sample Design, with the actual process of drawing the sample being the responsibility of Ipsos-MORI. In terms of the sample structure two key features of the panel have an influence:

- **Geography.** Some form of differential sampling by broadcast area is needed to account for differential co-operation rates and where panel sizes have been boosted. The Establishment Survey sample reflects this, resulting in a stratified sample design.
- **Clustering.** The extent of clustering in the panel design has a direct bearing on the clustering of Establishment Survey interviewer assignments.

1.2.2 Coverage

The domain of study is all private households (and permanent residents) in the United Kingdom, Channel Isles and Isle of Man.

1.2.3 Sampling Units

The sample will consist of a list of addresses allocated into assignments for issue to interviewers. The actual selection of these addresses will be an Ipsos-MORI task, carried out to RSMB's specification. To create workable assignments some form of clustering of addresses will be required, meaning that the sample selection must at first consist of selecting some geographical unit as a primary sampling unit, within which addresses can be sampled. The sampling unit that will be used will be Postcode. There are over 1.7million postcodes in the UK with an average of 15 residential addresses, although in reality this can vary significantly across postcode.

1.2.4 Stratification

Geographical stratifiers that will be incorporated in the design are:

- ITV Areas
- BBC Regions
- Sub-geographic breakdowns (e.g. Scottish Central Belt)
- DTT Reception Density Areas
- Population Density
- Geodemographic
- Cable Areas (No/Yes (<22.2% Penetration)/Yes (>=22.2%Penetration))

All of these are either important reporting areas, obvious discriminators of viewing and equipment ownership, or both. RSMB hold and maintain a postcode sector file containing these area definitions and the latest version of this file will be supplied to Ipsos-MORI. RSMB will constantly review the suitability of the above indicators and the requirement for any additional stratifiers.

Sampling areas are created by the interlacing of ITV Areas, BBC Regions and sub-geographic breakdowns, which is fully detailed in Appendix below. Based on the sampling areas, RSMB have specified the required total number of interviews in each of these sampling areas (see Appendix A). Appendix B (supplied by Ipsos-MORI) contains the average number of addresses needed per assignment in these sampling areas to give the required 12 interviews per assignment.

There are two procedures for the selection of Postcodes, one for urban/suburban areas and one for rural areas (about 15% of the population), the latter where a more clustered design is necessary to create a workable fieldwork task.

Urban and rural areas will be defined using postcode sector population density. The criteria for this will be defined such that rural areas still represent 15% of the population.

Urban/suburban areas

Within each sampling area postcodes are ordered within the following hierarchy:

- Cable Areas (No/Yes (<22.2% Penetration)/Yes (>=22.2%Penetration))
- DTT Reception Density Areas (Low/Medium/High)

This creates up to 9 reception sub-strata. If a sub-stratum has a population estimate smaller than the sampling interval, it is collapsed into the adjoining stratum. Postcodes are then sorted by ACORN code within strata, and selected systematically with probability proportionate to population, using a random start point and fixed interval procedure. We would propose to select 4 start postcodes per assignment in England, Wales, Northern Ireland and Scotland with each identifying a separate sub cluster.

Rural areas

Within each sampling area postcode sectors are ordered within the 9 reception sub-strata. They are then sorted within these strata by “wealth indicator” (averaged ACORN code for the constituent Postcode) and selected systematically with probability proportionate to population, using a random start point and fixed interval procedure. As above we would propose to select 4 start postcodes per assignment in England, Wales, Northern Ireland and Scotland with each identifying a separate sub cluster.

Within these sampling areas, postcodes should be ordered within the following hierarchy:

Cable Areas (No/Yes (<22.2% Penetration)/Yes (>=22.2%Penetration))
DTT Reception Density Areas (Low/Medium/High)

Within these strata, postcodes are then sorted by ACORN code and selected systematically with probability proportional to private household counts, using a random start point and fixed interval procedure.

If a sector has a limited number of postcodes or residential addresses then these should be joined to a nearby rural postal sector, iteratively if necessary, until expanded area can facilitate the sampling methodology.

Cable representation

In addition, in order to improve cable representation on a monthly basis, there is a small degree of balancing employed to equalise the distribution after sampling with respect to cable area, whilst still ensuring there are no adverse effects on other stratifiers.

1.2.5 Sample Sizes

The 49 sampling areas are the primary stratification and sample requirements within the areas are calculated with reference to ITV area panel sizes, historical co-operation rates and the following requirements:

The annual sample size is 53,004.

The design should be split into four quarterly replicates, requiring segment samples to be calculated in multiples of 48 to generate a whole number of assignments of 12 interviews per quarter.

The following will be indicative of annual sample sizes by BBC Region and ITV Panel Control areas as these represent the initial target interview at the start of the contract, although these can vary due to changes in co-operation rates.

ES Annual Sample Sizes by Area

BBC Regions

503	BBC East	2736
504	BBC West	1632
505	BBC South West	1680
506	BBC South	3264
507	BBC Yorkshire and Lincolnshire	3552
508	BBC North East and Cumbria	2748
509	BBC North West	5616
510	BBC Scotland	5232
511	BBC Ulster	3552
512	BBC Wales	3648
513	BBC Midlands West	4512
514	BBC Midlands East	2448
515	BBC London	10800
516	BBC South East	1584

ITV Panel Control Areas

1	London	11088
2	West Midlands	4608
3	East Midlands	2400
4	North West	5616
5	Yorkshire	3696
6	North East	2076
7	Scottish	3600
8	Grampian	1248
9	South	3216
10	South East	1392
11	East of England	2784
12	South West	1584
13	Ulster	3552
14	Border	960
15	West	1536
16	Wales	3648

1.2.6 Address Selection

For the core Establishment Survey of households the best sampling frame will be the small-user Postcode Address File (PAF), which is the closest approximation to a complete list of households available. This is the sampling frame used for the current BARB Establishment Survey and we would recommend its continued use.

Key issues regarding address selection are:

- The degree of clustering imposed (i.e. sampling interval).
- Differential response rates by area/geodemographic.
- Planning for ineligible addresses (e.g. business addresses).

The sample is selected from the PAF with the addresses located in geographically adjacent postcodes. The number of addresses issued varies by region, being

designed to yield an expected number of twelve interviews in a whole interviewer assignment. The number of target addresses across assignments would range between c.14 to 21 addresses. Within the same sub-cluster, the basic sample of addresses is matched by a substitute sample selected in the same way from a different random start point. Substitutes are used strictly in sequence as replacements for ineligible addresses in the basic sample. Once sampled, Postcodes are flagged and will not be sampled in the following two years. Also postcodes are excluded which have a current panel household.

In order to reflect panel design requirements, the Establishment Survey sample frame has to accommodate the following:

- Panel requirements
- Regional recruitment rates
- Regional dropout rates

1.2.7 Allocation of the Sample to Fieldwork Time Periods

The sample must be representative over time to ensure valid time-series data for projection of dynamic universes. The procedure will be to select the sample annually and systematically split this into quarterly replicates. These replicates will be representative of ITV areas and BBC Regions. Quarterly samples are then split into monthly samples representative of the Network, and the fieldwork operation uses these monthly samples to define the fieldwork task in terms of monthly waves.

For the yearly samples, assignments will be ordered within sampling area by the various reception sub-strata of the first postcode, and allocated to months in cyclical fashion as follows:

	Month
Assignment 1	1
Assignment 2	4
Assignment 3	7
Assignment 4	10
Assignment 5	2
Assignment 6	5
Assignment 7	8
Assignment 8	11
Assignment 9	3
Assignment 10	6
Assignment 11	9
Assignment 12	12
Assignment 13	1

....

And so on.

The current algorithm already developed to start each sampling area with a different month should continue.

In addition the process will optimise the distribution of sample points by cable stratifier in order to remove the worst of the monthly imbalances that this allocation process creates.

1.2.8 Multiple addresses at issued households

If two or three households are resident at an issued address, an interview is attempted with each one. If more than three addresses are found, a Kish grid procedure is employed to identify which three households should be approached.

1.2.9 Exclusion of Postcodes from sample

In order to reduce the possibility of clustering on the ES and the panel, the postcodes drawn are checked and substituted if any have been used in the previous 2 years' sample. The same process of checking and substituting is also applied for any postcodes containing existing panel homes.

1.3 Establishment Survey Fieldwork

The annual sample is selected in such a way that it can be systematically divided into twelve representative replicates as far as the targets permit. These are conducted as twelve consecutive four weekly surveys, with an additional two week overlap reissue period. The timetable is structured such that the Christmas/New Year break is missed.

Each monthly sample will support Network reporting and each quarterly sample will support full regional reporting.

The fieldwork is conducted in respondents' homes by trained interviewers using computer assisted personal interviewing (CAPI). Interviewers must attempt to interview at all of their issued addresses. In the case of an address being empty/demolished/not located (after checks), not a private address, not located or a holiday home the interviewer may use a substitute address. At least four calls must be made at each address, at least two of which must be on a weekday evening or at the weekend, before an address can be classified as a non-contact. The interview may be carried out with one of the following:

- the housewife or joint housewife
- the head of the household or joint head of household
- the spouse/partner of the housewife or head of household.

All assignments are subject to a strict policy of re-issue in the event of a low initial response, with a back-check on 10% of all assignments to ensure that correct procedures have been followed.

1.4 Establishment Survey Editing Procedures

Wherever possible inconsistencies or ambiguities are queried and corrected during the interview – the use of CAPI enables checks to be made while the interview is in progress. Two types of CAPI edit checks are used:

Hard checks

These checks, which do not allow the interviewer to enter an incorrect value, are used for answers each of which can only take a predefined value (e.g. 1 yes; 2 no; 8 don't know; 9 refused).

Soft checks

These are where the interviewer can, if necessary, override the check and enter the original answer (having double checked it with the respondent); these are used for plausibility checks and some range checks (e.g. a claimed telephone number not equating to 11 digits would have to be queried with the respondent before the number could be entered).

Questionnaire information is returned by the interviewer each evening and is initially processed by an automated validation procedure to ensure that all interviews have been completed correctly. Any interviews that have not successfully been processed by the automated system are manually checked against a Field Management System to ensure that the serial numbers of addresses interviewed at are validated against a second source.

Further more stringent checks of the data are carried out once the interview has been validated. These include all social grades which are reviewed in the office with any miscodes identified and updated.

Additionally, particular attention is paid to checking the accuracy of key questions relating to the reception equipment available to the household. This includes:

Platform reception
ITV regional reception
Free-to-air or subscription-based reception for satellite

A complex editing procedure is employed to ensure that this information is coded correctly and non-response is minimised.

1.5 Establishment Survey Weighting

1.5.1 Overview

The weighting involves weighting households and individuals to Government-derived population estimates such that:

- A separate weight is generated for monthly, quarterly, and rolling four-quarterly. This results in each interview being assigned six processing weights:

Monthly
Quarterly
Annual when interview is in first of four rolling quarters
Annual when interview is in second of four rolling quarters etc....

- In each of these cases, the weight produced for each interview must satisfy the overall grossing-up requirements of households and individuals simultaneously. It is unacceptable for individuals within a home to carry different weights.
- The weighting controls for households and single person households by BBC/ITV segment, S4C segment and for sex interlaced with age by Standard Region. To achieve this it is necessary to use rim weighting.
- S4C boost interviews are included in the weighting and subsequent analyses.

1.5.2 The Rim Configuration

There are four rims; Rims 1, 2 and 3 are used for quarterly and annual weighting. Rims 3 and 4 are used for monthly weighting.

Rim 1 Sampling area by HH Size 1/ HH Size 2+. This has 98 cells (49 Sampling areas by 2 HH Size splits).

Rim 2 Households by S4C Segment. This has 11 cells (10 S4C segments in Wales plus outside Wales).

Rim 3 Sex by Age by Government Office Region. This has 216 cells as follows:

Sex (Male/Female) by
 Age (0-4,5-9,10-14,15-24,25-34,35-44,45-54,55-64,65+)
 Government Office Region (London, South East, South West, West Midlands, East Midlands, East Anglia, North West, Yorkshire & Humberside, North, Wales, Scotland, Northern Ireland). Note that Channel Isles is included in South West, Isle of Man in North West.

Rim 4 ITV Panel Control Area by HH Size 1/ HH Size 2+. This has 32 cells – 2 HH Size by 16 ITV areas:

London
 West Midlands
 East Midlands
 North West
 Yorkshire
 North East
 Scottish
 Grampian
 South
 South East
 East Of England
 South West
 Ulster
 Border
 West
 Wales

1.5.3 Targets (Population projections)

It is important to ensure that the appropriate month's targets are employed to generate the appropriate weight for each period. The principle is to use the latest month's targets, whatever the period. This is illustrated below, using August 2014 data as an example.

<u>Application</u>	<u>Use targets from period</u>
Monthly (Aug 2014)	Aug 2014
Quarterly (Jul-Sep 2014)	Sep 2014
Annual with Aug 2014 in 1st of four quarters (Jul 2014 - Jun 2015)	Jun 2015
Annual with Aug 2014 in 2nd of four quarters (Apr 2014 - Mar 2015)	Mar 2015
Annual with Aug 2014 in 3rd of four quarters (Jan 2014 - Dec 2015)	Dec 2015
Annual with Aug 2014 in 4th of four quarters (Oct 2013 - Sep 2014)	Sep 2014

1.5.4 Weighting to Household and Individual targets simultaneously

Rim 3 differs from the rest in having individual targets, not household targets. However the weighting still operates on households, not individuals.

1.5.5 Convergence and number of iterations

The targets are in thousands. The convergence criterion ensures that each target is met exactly after rounding.

1.5.6 Collapsing of cells

For a single quarter, one of the household size splits in a small sampling area may have no interviews, preventing convergence in Rim 1. In such a case, the two household size cells are collapsed into a single total households cell for the segment.

In addition there may be instances where very small sampling areas have no interviews at all. The principle here is to collapse cells in a geographical hierarchy as follows.

Sampling area

Collapse sampling area into BBC/ITV Segment. For example, if there are no interviews in Sampling Segment 3.1 collapse 3.1, 3.2 and 3.3 into BBC/ITV Segment 3.

BBC/ITV Segments

Collapse BBC Regions. For example, if there are no interviews in Segment 20 (BBC West-ITV South) then collapse with segment 21 (BBC South-ITV South).

S4C Segment

First, collapse Welsh Speaking Dimension (H,M,L), i.e. preserve North/West/South and Core/Overlap.

Second, if necessary, collapse Core and Overlap, i.e. preserve North/West/South.

APPENDIX - Sampling Area Segment Definition

Segment	ITV Area	BBC Region Definition	Split Definition
Seg 1	London	East of England	No London split
Seg 2	London	South	No London split
Seg 3	London	London	Inner London
Seg 4	London	London	Outer London
Seg 5	London	London	Remainder
Seg 6	London	South East	No London split
Seg 7	West Midlands	East of England	Not Met C
Seg 8	West Midlands	West	Not Met C
Seg 9*	West Midlands	South	Not Met C
Seg 10	West Midlands	Yorkshire and Lincolnshire	Not Met C
Seg 11	West Midlands	Midlands West	MC: West Midlands
Seg 12	West Midlands	Midlands West	Not Met C
Seg 13	East Midlands	Midlands East	Not Met C
Seg 14	North West	Yorkshire and Lincolnshire	Not Met C
Seg 15	North West	North West	MC: Greater Manchester
Seg 16	North West	North West	MC: Merseyside
Seg 17	North West	North West	Not Met C
Seg 18	North West	Midlands West	Not Met C
Seg 19	Yorkshire	Yorkshire and Lincolnshire	MC: South Yorkshire
Seg 20	Yorkshire	Yorkshire and Lincolnshire	MC: West Yorkshire
Seg 21	Yorkshire	Yorkshire and Lincolnshire	Not Met C
Seg 22	Yorkshire	North East and Cumbria	Not Met C
Seg 23	Yorkshire	Midlands East	Not Met C
Seg 24	North East	North East and Cumbria	MC: Tyne & Wear
Seg 25	North East	North East and Cumbria	Not Met C
Seg 26	Scottish	Scotland	Scottish central belt
Seg 27	Scottish	Scotland	Not Scottish central belt
Seg 28	Grampian	Scotland	No Split
Seg 29	South	West	No Split
Seg 30	South	South	Exclude Channel Island
Seg 31	South	South West	Channel Island
Seg 32	South	South East	No Split
Seg 33	South East	South East	No Split
Seg 34	East Of England	East of England	No Split
Seg 35	East Of England	London	No Split
Seg 36	South West	South West	No Split
Seg 37	Ulster	Ulster	Ulster Geography 1
Seg 38	Ulster	Ulster	Ulster Geography 2
Seg 39	Ulster	Ulster	Ulster Geography 3
Seg 40	Ulster	Ulster	Ulster Geography 4
Seg 41	Ulster	Ulster	Ulster Geography 5
Seg 42	Border	North East and Cumbria	No Split
Seg 43	Border	North West	No Split
Seg 44	Border	Scotland	No Split
Seg 45	West	West	No Split
Seg 46	West	South	No Split
Seg 47	Wales	Wales	S4C North
Seg 48	Wales	Wales	S4C West
Seg 49	Wales	Wales	S4C South
Seg 50	Wales	Wales	S4C Cardiff Area

**Segment 9 no longer exists following Area definition change.*

London Geography: London Split by inner/outer/old definition Outer & Remainder. **Metropolitan**

County Definition: Geographical Regions defined by JICPOPS

Scottish Central Belt: Defined by Local authority (highest population density)

Ulster Geography: Split by % of population who are Catholic (0-20%, 21-40%)

S4C: Defined by S4C Segment