



SHS - design and implementation

Steven Hope

Deputy Managing Director

Ipsos MORI Scotland

Sampling

- Driven by clear sampling objectives
 - Requirements at different geographical levels
 - Requirements at different points in time
- Preference for unclustered design
- Most cost-effective design
 - Cost per effective interview - total fieldwork cost divided by effective sample size
 - Long term desire to reduce clustering

Sampling objectives

- Expressed in terms of simple random sample equivalent of
 - 2,500 interviews in Scotland as a whole for each quarter in each year
 - 500 interviews in each local authority with over 120,000 households in each year (the Big Five local authorities – Glasgow, Edinburgh, North Lanarkshire, South Lanarkshire and Fife)
 - 500 interviews in each local authority (regardless of size) in each two-year period
 - 500 interviews in each category of the six-fold urban rural classification in each year, and
 - 2,000 interviews in the 15% most deprived areas of Scotland, taken together as a group, in each year.
- Design needs to take account of complex standard errors associated with selection probability, corrective weighting and clustering

Sample design

- Disproportionate stratification by LA
 - Minimum to meet SRS requirement after two years, otherwise proportionate
- Combination of clustered and unclustered sampling
 - Across local authorities – each LA split into urban and non-urban strata
 - 16 LAs wholly unclustered after 2 years
 - 2 LAs wholly clustered
 - Overall, 70% of sample unclustered after 2 years
- In general, urban areas use unclustered sampling and non-urban use clustering but ...
 - Western Isles, Orkney, Shetland and Clackmannanshire use unclustered samples
 - Minimum sample requirement and small population means sampling fractions of 1 in 8, 7, 6 and 17 respectively.
- Allocating sample to achieve savings from clustering at intermediate time periods

Sample allocation

- Batching
 - Organising sample into the most efficient interviewer allocations
 - Minimise distance between addresses
 - Lower travel costs
 - More contact time – less travel time
 - More calls at addresses
- Clustered areas
 - Primary sampling units (data zones) form natural clusters. Average sampling fraction of about 1 in 20
- Unclustered areas
 - Large LAs – addresses assigned to survey years and then batched. Gives clustered sample in each quarter, unclustered each year. Average sampling fraction about 1 in 100
 - Other areas – addresses batched then batches assigned to survey years. Gives clustered sample each quarter AND each year but unclustered after two years. Average sampling fraction about 1 in 55

Questionnaire design

- Three broad priority areas
 - Communities
 - Transport
 - Local government
- Consistent core questions
 - Consistent demographic and topic core
 - Major changes at start of each two-year sampling period
 - Minor changes in intermediate years
- Streaming
 - Following review of SHS, a more core/module approach adopted in 2007 to help manage demand
 - Each sampled address allocated to one of ten 'streams'
 - Proportion of sample asked each question controlled by CAPI routing

Data collection and management

- Data collected by CAPI
 - 1999-2006 used In2itive
 - 2007 and 2008 combination of In2itive and Blaise
 - Major CAPI problems in 2007 – scale and complexity of survey script breached programming limitations of In2itive
 - January 2009 – TNS switched to NIPO
 - July 2009 – Ipsos MORI switched to Quancept
- SHS only possible as CAPI
 - Paper version of questionnaire over 400 pages
 - Complex routing and look-up tables
 - Travel diary
 - Randomisation
 - Selection of random adult, random child, random school child
- Data processing by contractors
 - Ipsos MORI – main data
 - TNS – travel diary