# USING CENSUS OUTPUT AREAS FOR MARKET RESEARCH SAMPLING

## Martin van Staveren

**Small Area Geography – Planning for the Future**

**MRS CGG Seminar,   5th February 2007**

# INITIAL USE OF GEODEMOGRAPHICS

- BMRB was first company to link Census geodemographics to market research data - 1979

- TGI data showed stronger relationship between product/brand usage and geodemographics than for social grade

- So geodemographics could be used as a powerful discriminator in sample area selection

- From 1983, used a master sample of Enumeration Districts (EDs) selected by region and ACORN type

- Samples for each survey selected to a matrix, representative by region and ACORN type

- Matched samples could be selected for Pre/Post or Continuous surveys

MRS.

# INSITE SAMPLING SYSTEM

- Computerised system introduced in 1990

- Database of ALL Census EDs, linked to demographic data, ACORN type and PAF listing

- Much greater flexibility than partial, paper system

- Automated selection of samples to specified parameters

- For continuous or repeat jobs, could select points to match a specified matrix

- Could track EDs already used, and re-use in a controlled manner

- Easy to select similar adjacent areas (additional or replacement) if required

# CENSUS AREAS – 1981 & 1991

Census statistics released for EDs

- EDs were designed before Census, for operational field workloads

- So size and shape suitable for face-to-face interviewers

- Average size about 150 households, but BMRB only used EDs with 90+ addresses, increasing average size to 200+ addresses

- EDs varied considerably in size and content – not designed to be statistically homogenous areas

- Boundaries not available, so census agencies derived approximate address list from centroids

# CENSUS AREAS - 2001

Census statistics for Output Areas (OAs)

- OAs designed after Census, for statistical purposes, so much more homogeneous

- Typical size for England and Wales is 125 households, with much less variation

- OAs built from unit postcodes & nest within wards

- But smaller size and falling response rates make them less suitable for survey research

- Not designed for interviewers – less regular shapes

- So need to combine OAs to form new Sample Units

MRS.

# Sample Units - the Brief

Need continuity & homogeneity

- Areas larger than previous EDs preferable

- Decided on one-off grouping exercise, in preference to matched samples for each survey:

  - Need to group Scotland as in 1991

  - Matching will be more difficult over time

  - Cost

- Basic plan was to merge adjacent OAs within a ward to give average SU of 250-300 households – slightly larger than average issued ED size

- Also take account of the OA level ACORN solution

MRS.

# Sample Units - The Process

Some OAs (over 200 households) large enough to use as Sample Units in their own right

- Identify remaining OAs within each ward that share a boundary

- Rank available merges by ACORN type difference and significance of shared boundary

- Merge OAs with only 1 neighbour with same ACORN type

- Merge other OAs of same ACORN type

- Merge OAs of different ACORN types starting with ±1type

- Merge remaining OAs with smallest available unit – usually a Sample Unit and over 200 households

- Manually edit large Sample Units  where shape impractical or where original merge not ideal

MRS.

# The Ideal

Example of a Bridge

A Bridge Too Far?

**Example in Scotland**
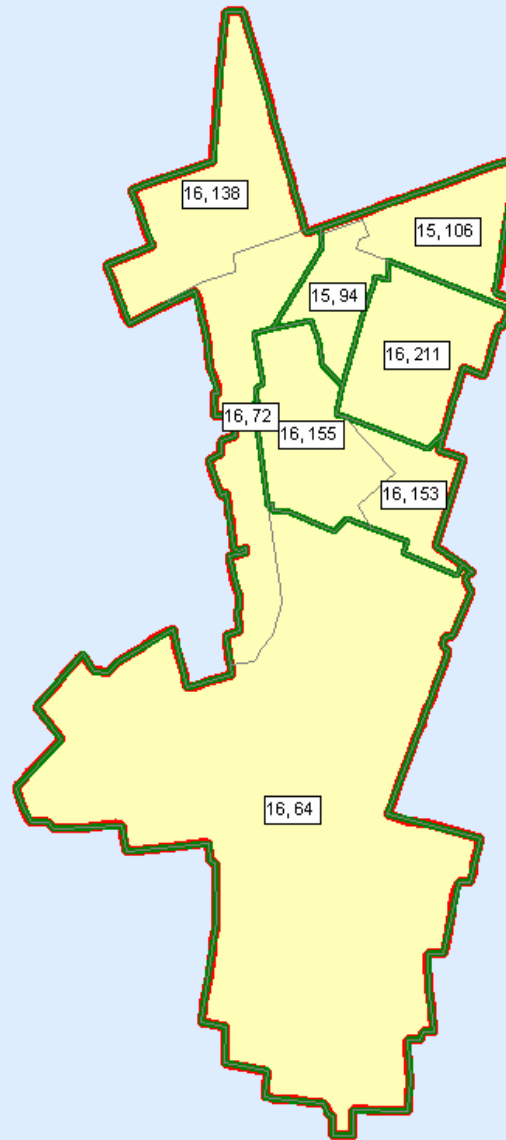
Legend:
- 2001 Wards
- Sample Units (from 2001 OAs)
- 2001 Output Areas
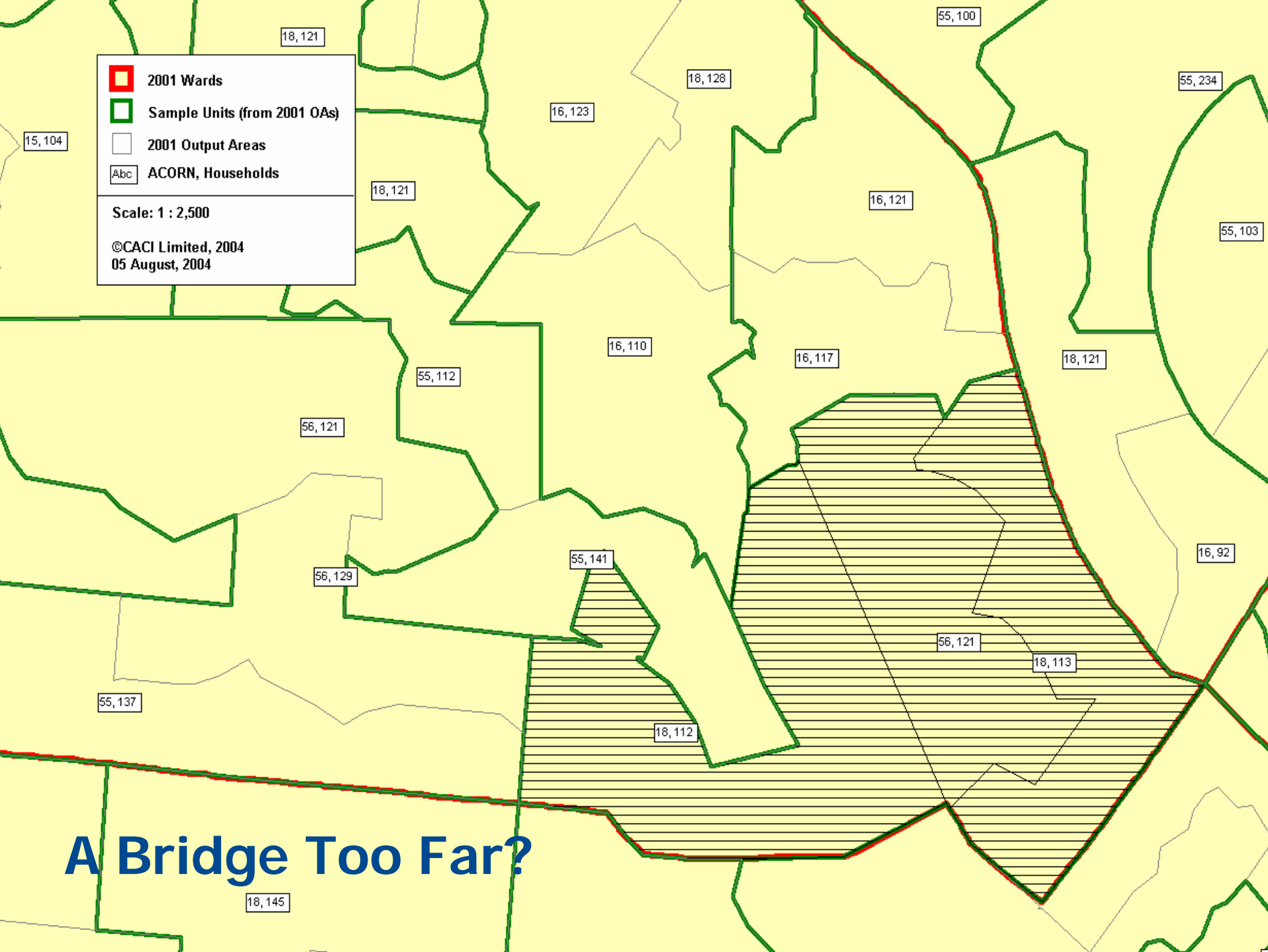- Abc  ACORN, Households

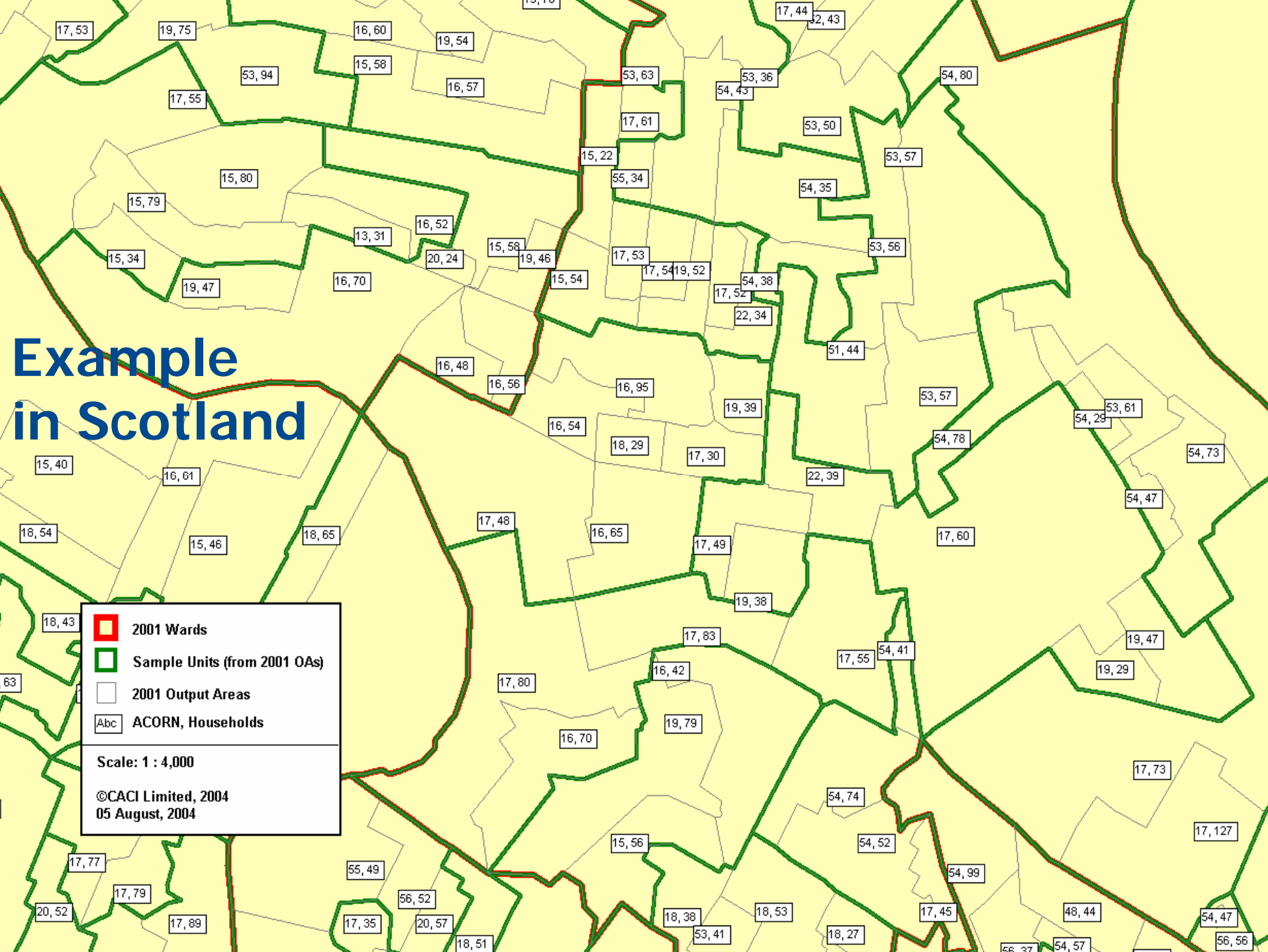Scale: 1 : 4,000

©CACI Limited, 2004
05 August, 2004

# Results

We now have 83108 Sample Units

- We have ensured continuity in our sampling processes while achieving larger areas built from more robust building blocks

- Average size is 300 households

- 62% of OAs are in Sample Units of the same dominant ACORN type: 66% in England & Wales, 48% in Scotland

- Now use more of the population - virtually all areas included (0.1% of pop. not in use against 2.9% of pop. using EDs)

# Distribution by Size
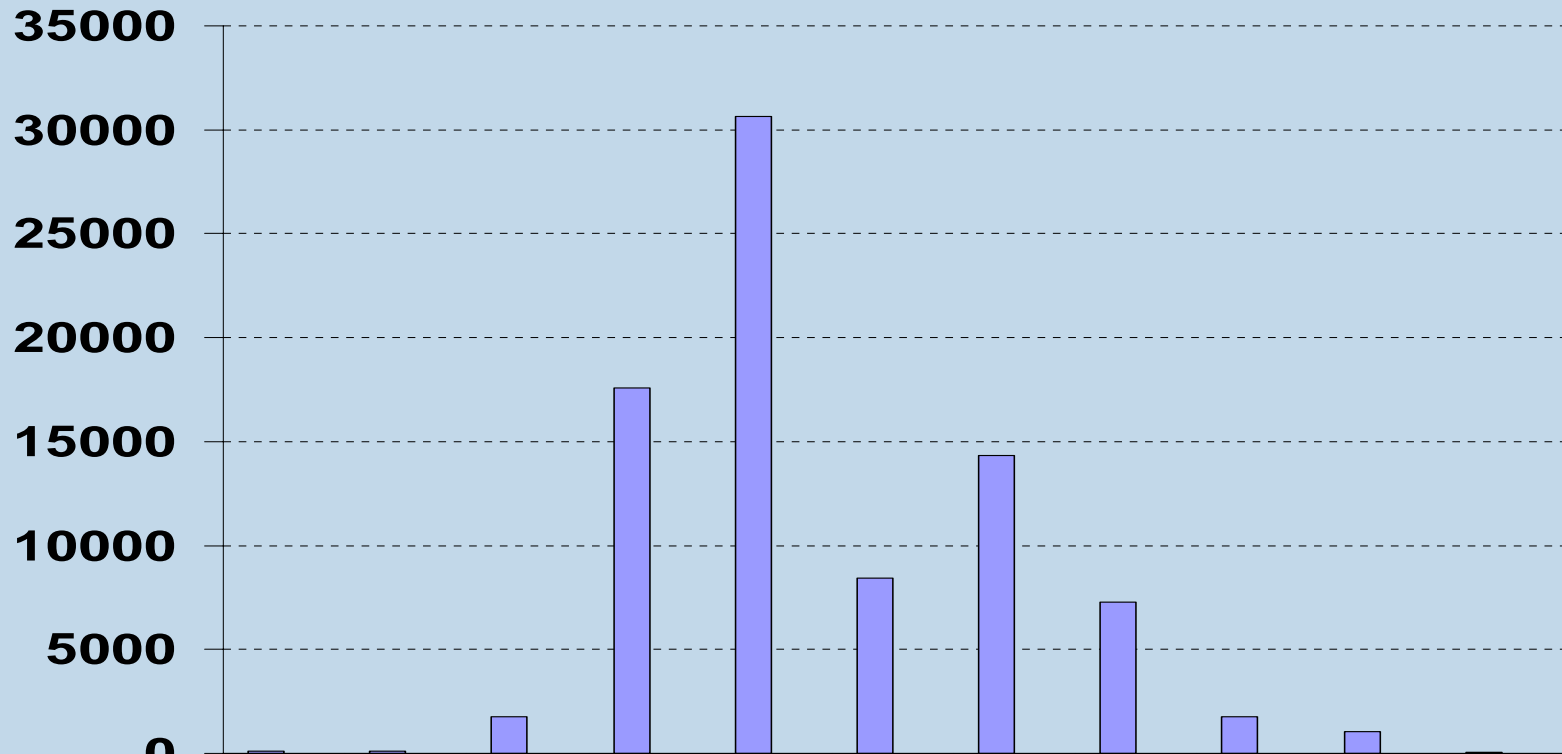


| | 0-99 | 100-149 | 150-199 | 200-249 | 250-299 | 300-349 | 350-399 | 400-449 | 450-499 | 500-999 | 1000+ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Count | 126 | 134 | 1773 | 1760 | 3065 | 8407 | 1431 | 7255 | 1770 | 1044 | 28 |
| Percent | 0.2% | 0.2% | 2.1% | 21.2 | 36.9 | 10.1 | 17.2 | 8.7% | 2.1% | 1.3% | 0.0% |

MRS.

# Current Position

Larger, more suitable areas

- A customised geography that is still compatible with Census areas

- Retained homogeneity from OAs - as far as possible

- Solution is durable for several years

- Continuity with past procedures

MRS.

# The future

- Declining use of Face-to-face interviewing in market research, but still the best method for some studies

- Areas of around 300 households suit our methods best – perhaps a little larger if response rates continue to fall

- Preference for areas which nest within wards and Local Authorities, so sampling points can be selected with reference to larger geographies

- Homogenous areas make sampling statistically more efficient – less variation within areas means a lower Design Effect

MRS.

# The future

Adjust areas to retain homogeneity, size & compatibility with LA geography

v.

Maintain areas to be comparable with 2001 **X**

Can cope with a one-off change in boundaries if the consequence is better homogeneity and more uniform size.

MRS.