

USING CENSUS DATA FOR MARKET RESEARCH SAMPLING

Martin van Staveren

&

Chris Walker

INITIAL USE OF GEODEMOGRAPHICS

- BMRB was first company to link Census geodemographics to market research data (Ken Baker 1979)
- TGI sample points classified by ACORN at ward level, post-hoc
- TGI data showed stronger relationship with product/brand usage for geodemographics than for social grade
- So geodemographics could be used as a powerful discriminator in sample area selection

ED MASTER SAMPLE

- Master sample of Enumeration Districts (EDs) selected by region and ACORN type, used from 1983 (paper-based)
- 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
- Could also use for samples restricted to certain Regions and/or ACORN types

ROLLING ACORN SAMPLES

- Developed in 1985
- For greater matching between waves of a continuous tracking study
- Sample for each wave consists of:
 - Even numbered addresses from one set of EDs, also used in previous wave
 - Odd numbered addresses from another matched set of EDs, also to be used in next wave
- So each wave has a 50% match with previous wave, and a 50% match with the following wave

INSITE SAMPLING SYSTEM (i)

- Computerised system introduced in 1990
- Database of all Census EDs, linked to demographic data, ACORN type and PAF listing
- All EDs available – not just a master sample
- Much greater flexibility than partial, paper-based system

INSITE SAMPLING SYSTEM (ii)

- Automated macros allowed samples to be selected to specified parameters
- For continuous or repeat jobs, can 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 (i)

- For 1981 and 1991, Census statistics were released for Enumeration Districts
- EDs were designed by OPCS pre-Census, for operational field workloads
- So size and shape suitable for face-to-face market research interviewers
- Average size about 150 households, but BMRB only used EDs with 90+ addresses, increasing average size to 200+ addresses

CENSUS AREAS (ii)

- EDs varied considerably in size – from 20 to 2000 addresses
- Boundaries were not available from OPCS, so census agencies derived approximate address list from ED centroid
- Were also variable in content – not designed to be statistically homogenous areas

CENSUS AREAS (iii)

- For 2001, Census statistics have been released for Output Areas (OAs)
- OAs designed post-Census, for statistical purposes, so much more homogeneous
- Typical size for England and Wales is 125 households, with much less variation around this figure
- OAs built from unit postcodes and nest within wards

CENSUS AREAS (iv)

- But smaller size and falling response rates make them less suitable for survey research
- Geography also not designed for interviewers – less regular shapes
- CACI / BMRB have combined OAs to form new Sample Units
 - Typical size is now 300 households
 - Try to retain homogeneity as much as possible

The Brief (i)

- Need for continuity – sampling routines within Insite, deliverables
- Areas larger than EDs preferable
- OAs smaller than EDs although larger than initially planned, in England and Wales at least
- Scotland a perennial problem

The Brief (ii)

- Alternative solutions identified as:-
 1. Matched samples for each study, as required
 2. One-off grouping exercise

- Decision to group based on
 1. Need to group Scotland as in 1991
 2. Matching would become more difficult over time
 3. Cost

The Brief (iii)

- The basic plan was to merge adjacent OAs within a ward to give an average SU size of 250-300 HHs – slightly larger than the average ED size we knew we were issuing
- Tim Temmink at CACI refined this to take account of the OA level ACORN solution that already existed

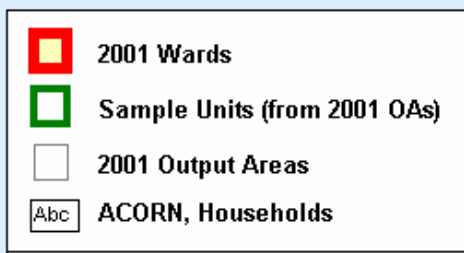
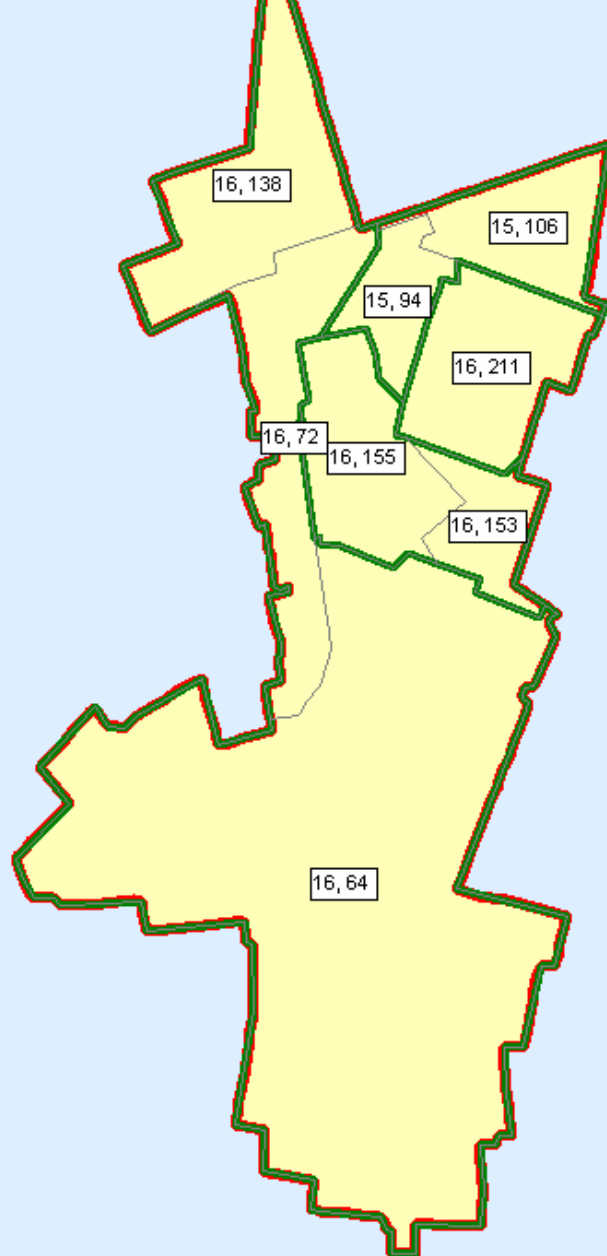
The Process (i)

1. Save large enough OAs (over 200HHs) as Sample Units in their own right
2. Identify OAs within each ward that share a boundary
3. Rank available merges by ACORN type difference and significance of shared boundary
4. Merge OAs with only 1 neighbour and with same ACORN type
5. Merge other OAs of same ACORN type
6. Merge OAs with different ACORN types starting with +/-1 type

The Process (ii)

7. Merge remaining OAs with smallest available unit – usually a Sample Unit and over 200HHs
8. Merge small Sample Units – exclusively in Scotland
9. Manually edit large Sample Units where shape impractical or where original merge not ideal

This involved roughly 17 different stages and numerous iterations within each stage

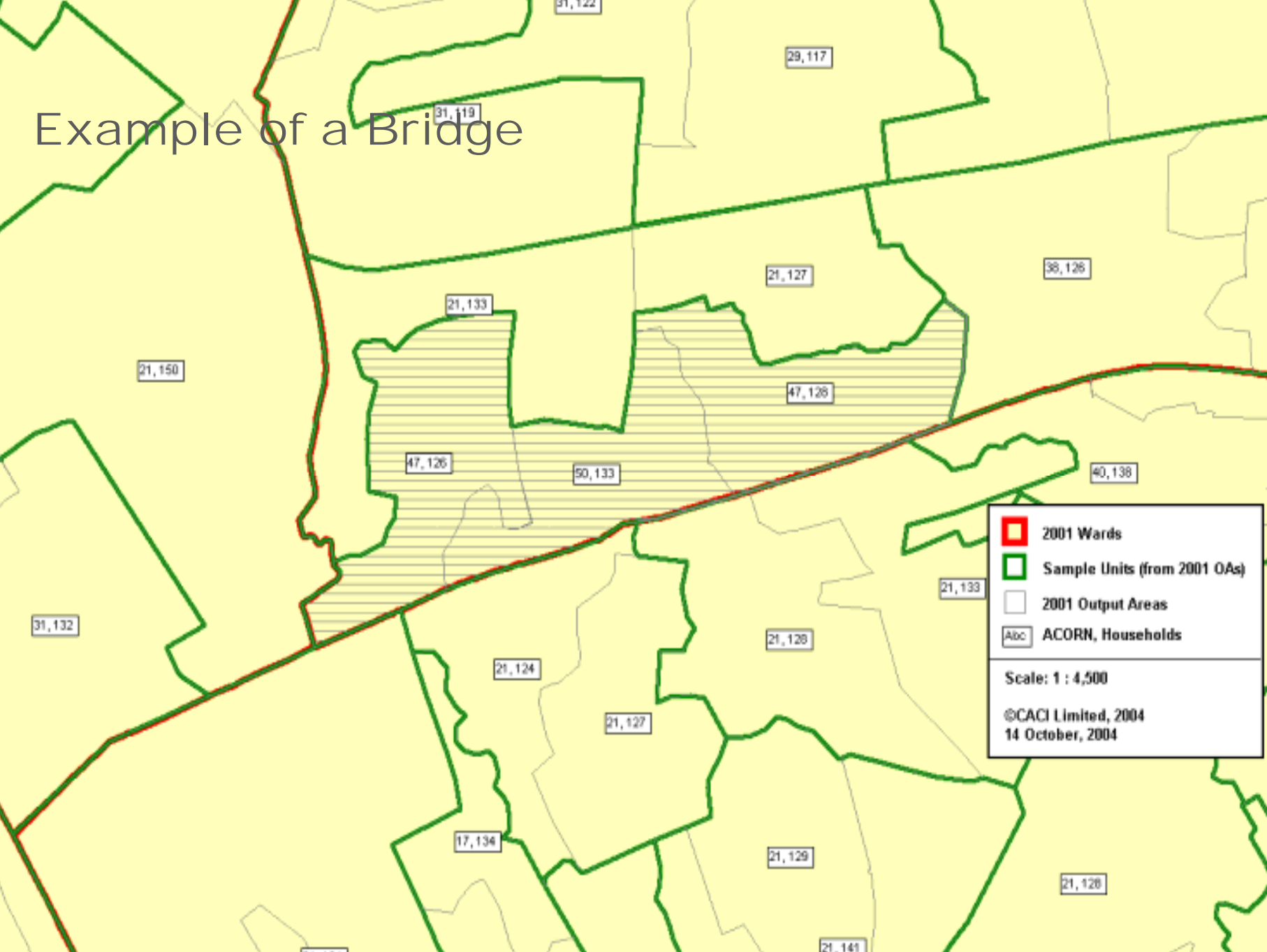






Scale: 1 : 4,622

©CACI Limited, 2004

05 August, 2004

Example of a Bridge

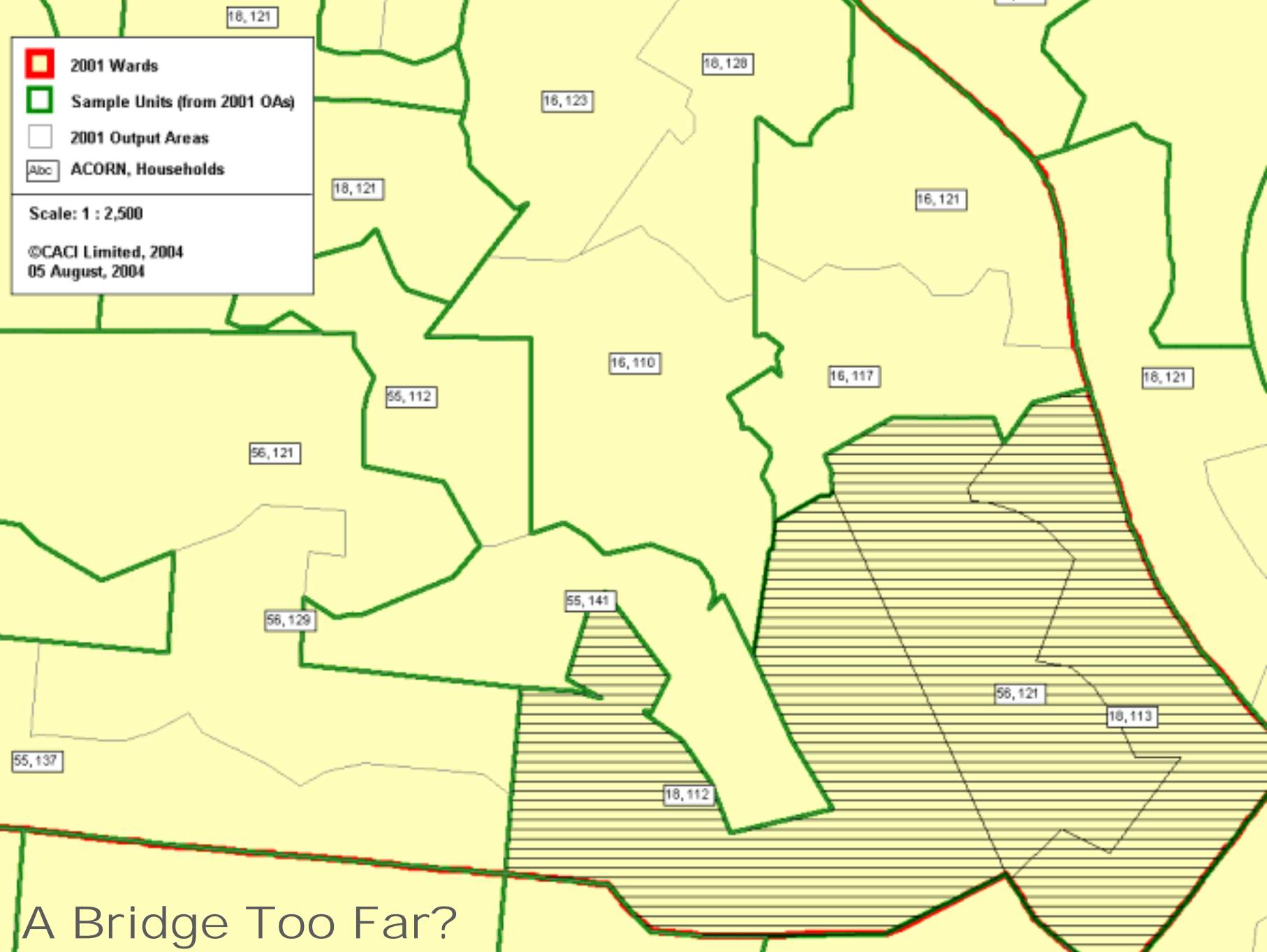


	2001 Wards
	Sample Units (from 2001 OAs)
	2001 Output Areas
	ACORN, Households
Scale: 1 : 4,500	
©CACI Limited, 2004 14 October, 2004	

- 2001 Wards
- Sample Units (from 2001 OAs)
- 2001 Output Areas
- ACORN, Households

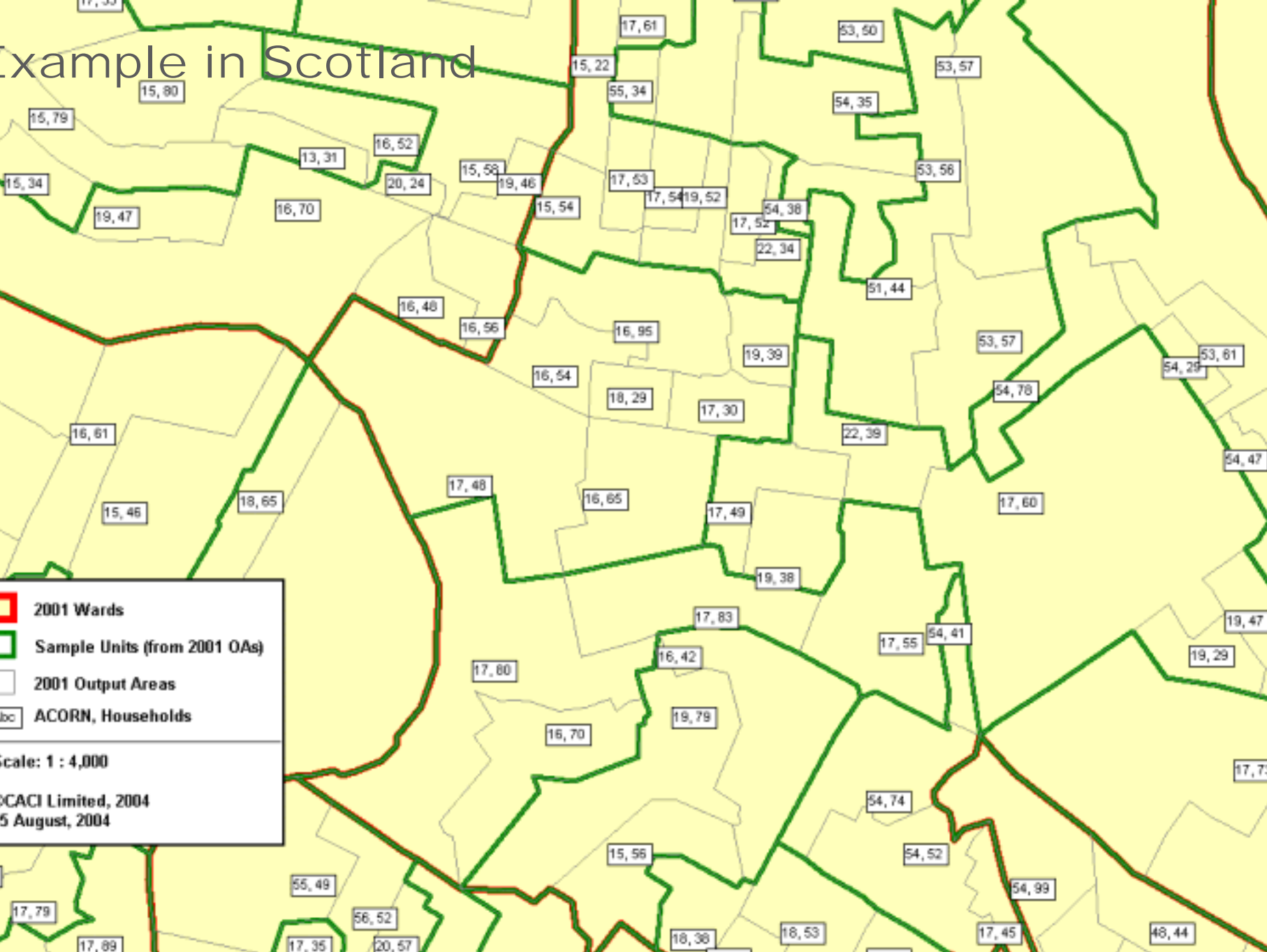
Scale: 1 : 2,500

©CACI Limited, 2004
05 August, 2004



A Bridge Too Far?

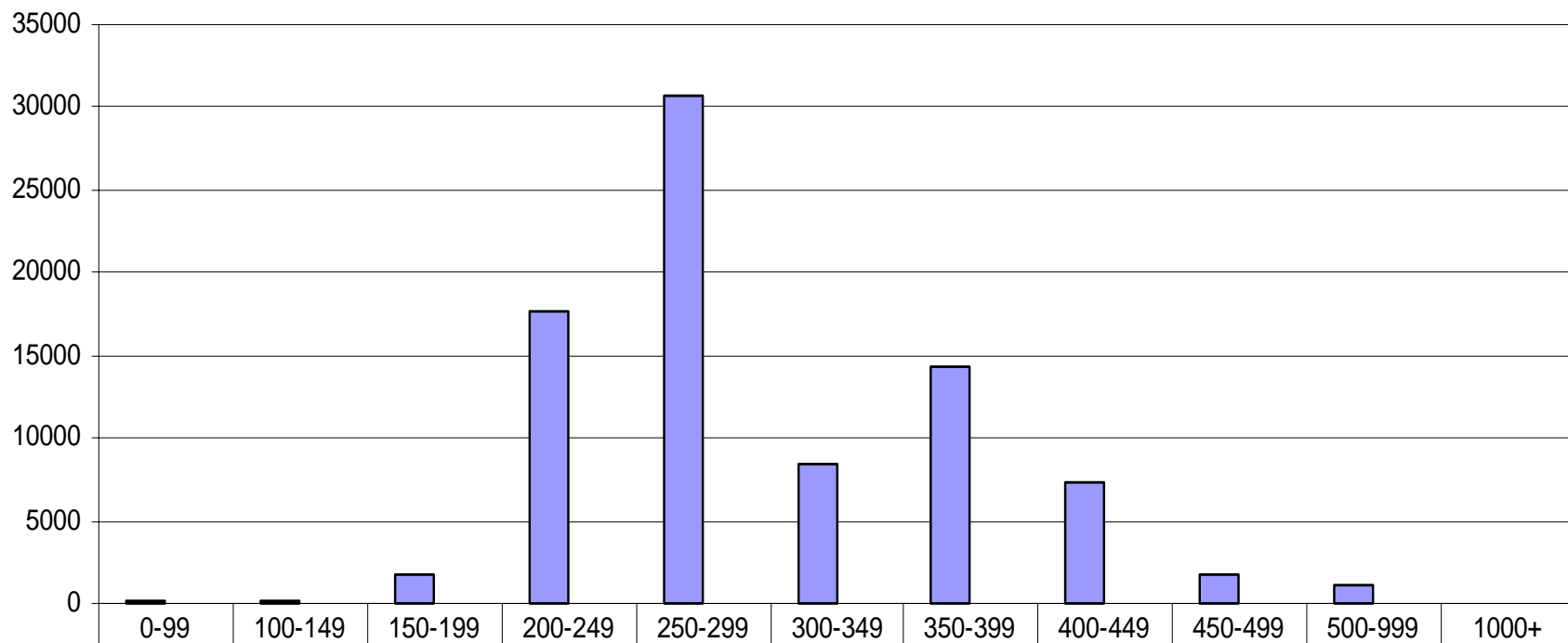
Example in Scotland



Results

- We now have 83,108 Sample Units
- We have ensured continuity in our sampling processes while achieving larger areas built from more robust building blocks
- Average size is 300 HHs
- 62% of OAs are in SUs 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 v 2.9% of pop. using EDs)

Distribution by Size



Count	126	134	1773	17601	30658	8407	14312	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%

Conclusion

- Larger areas - more suitable for Field
- We have 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