



National Statistics Small Area Geography Policy (England and Wales)

- Review and Consultation

1	INTRODUCTION.....	2
2	REVIEWING OUTPUT AREAS AND SUPER OUTPUT AREAS.....	3
3	THE PROPOSAL.....	4
4	SMALL AREA GEOGRAPHY POLICY CONSULTATION QUESTIONNAIRE....	7

APPENDICES

5	BACKGROUND.....	19
5.1	2001 CENSUS – OUTPUT AREAS.....	19
5.2	NEIGHBOURHOOD STATISTICS – SUPER OUTPUT AREAS	20
5.3	IMPLEMENTING PAT 18 REPORT RECOMMENDATIONS AND THE OA / SOA HIERARCHY	21
6	TOPIC NOTES.....	26
6.1	STABILITY OR RE-DRAWING	26
6.2	SCALE OF REAL WORLD MAINTENANCE	28
6.3	METHOD OF APPLYING MAINTENANCE FOR REAL WORLD CHANGES	30
6.4	ALIGNMENT TO LOCAL AUTHORITY BOUNDARIES.....	31
6.5	STRUCTURAL MAINTENANCE	31
6.6	LINKS TO POSTCODES.....	31
6.7	COMMUNAL ESTABLISHMENTS.....	32
6.8	SDC METHODS AND OVERLAPPING GEOGRAPHIES	32
6.9	BOUNDARIES AND LICENSING.....	33
6.10	OA / SOA CONTIGUITY AND ZERO POPULATION OAS.....	34
6.11	UPPER LAYER SOAS.....	34
6.12	BUSINESS OUTPUT AREAS.....	35
6.13	NAMING AND CODING.....	35
7	CONTACT DETAILS	37

1 Introduction

Output Areas (OAs) were developed as the geography for reporting small area statistics from the 2001 Census. Super Output Areas (SOAs) were proposed in 2003 as a layered geography for Neighbourhood Statistics. The Lower and Middle Layers of an intended three layered hierarchy were released in 2004 (LSOAs and MSOAs). An Upper Layer is yet to be constructed (USOAs).

This paper provides an opportunity for you to contribute to the future National Statistics small area geography policy for England and Wales. In the autumn of 2005 a proposed policy was drafted and tested with Census Advisory Groups. Its main thrust was to maintain a high degree of stability with OAs and SOAs.

Stability was not envisaged when OAs were developed, but it was a key aim behind the creation of SOAs. There has never before been a non subjective geography specifically for small area statistical outputs. It enables data to be consistently compared and analysed over time and space. Where accurately geo-referenced low level data is made available to ONS, it brings an opportunity to output data for other user defined geographies. The potential the geography brings is clear where there has been cogent data to apply to it. For example, the Index of Multiple Deprivation (2004), produced against LSOAs, has had a tremendous impact for many. Stability with OAs also brings an opportunity, for the first time, to have a consistent small area geography between censuses.

ONS are now in the middle of a period of research to consider aspects of this proposed policy, and alternatives to it. As part of the research programme, the core proposal and key issues that arise, are now put forward for wider consultation. This consultation will run for 12 weeks. As well as consulting on the proposal, this is also an opportunity to review OAs and SOAs, their usage and usefulness during the last 2 to 3 years.

This paper is intended to seek the views of all stakeholders of the Census and Neighbourhood Statistics, together with users / suppliers of other small area datasets and analyses. Specific questions for consultation are included in section 4 of this paper, but comments on any aspect of ONS' small area geography policy are strongly welcomed. Please do take the time to respond to the consultation, whether you broadly support the policy outlined or not.

Responses to the consultation can be completed on-line at:

www.onsgeography.net

This site will also host a separate forum or 'blog' where issues will be posted and debate is actively encouraged. Alternatively, responses can be returned via e-mail or post to the contact points given in section 7. Please use the questionnaire template included in this paper when responding in this way. The consultation will end on the 16th February 2007. All responses should be returned by this date and will be made public.

2 Reviewing Output Areas and Super Output Areas

When SOAs were launched in 2004, a future review of them in 2006 was announced. This consultation is the opportunity for users and stakeholders to feed their comments into that review. The scope of the review has been broadened to also include OAs.

We are interested in understanding how OAs and SOAs have been used, and users' experience in using them. As part of this we are keen to receive feedback on OA / SOA usefulness in terms of:

- reporting and analysing 2001 Census results;
- reporting and analysing other datasets;
- using the Neighbourhood Statistics service;
- building other geographies of interest;
- matching different datasets;
- taking action as a result of data analysis.

We are also keen to understand if there are specific local examples of OAs or SOAs that have proved problematic to work with. If so, details of these are sought.

Finally, there may be potential uses identified for a stable small area geography that have not been able to be realised. Details of these are also sought.

3 The proposal

The proposal for small area geographies put forward to Census Advisory Groups at the end of 2005 was based on the following key principles:

- 2011 Census and Neighbourhood Statistics geography policies must be aligned;
- We need a primary output geography which is good for statistics and which can also enable policy development and analysis;
- OAs and SOAs are the fundamental geography for Neighbourhood Statistics and provide a consistent small area geography for time series analyses between censuses. As a result, the stability of the current OA / SOA hierarchy carries a very high weight;
- The potential for flexible output geographies and the release of past and future data against them still exists. To be realised it requires that unit records be accurately geo-referenced, made available to ONS where they are not collected by them, and that enabling disclosure control policies and methods be established.

These principles remain, accepting that as with most requirements and design criteria, a balance often has to be struck. With these as a basis ONS now proposes the following geography policy for small area outputs. It is largely the same as that put forward at the end of 2005, but incorporates a few points of important clarification.

1. The OA / SOA hierarchy will continue in its present form beyond the next Census.
[See sections 5 and 6.1]
2. “Definitive” 2011 Census outputs will be released for this hierarchy (OA – SOA – Local Authority).
3. Over time, ward boundaries may no longer align to OA or SOA boundaries. This degradation will continue. Wards will no longer be the main output geography. If there is a significant user demand, work will be undertaken to assess the potential for developing statistical techniques to produce census outputs and other key datasets for wards or alternative geographies.
[See section 6.8]
4. OA minimum size disclosure thresholds for resident population and households will remain at 100 and 40 respectively.
5. Maintenance of the hierarchy will be required to respond to changes in the real world and may be undertaken to fix locally identified existing problems. Overall, however, maintenance changes are expected to affect no more than 5% of OAs and SOAs.
[See sections 6.2 and 6.3]

5a. Where minimum population and household thresholds are breached, OAs and SOAs will be merged with a neighbour, enabling a simple lookup to be maintained between 2001 and 2011 areas. Research to date suggests that < 0.3% of OAs and SOAs will need to be merged for this purpose.

5b. Guideline upper thresholds may be introduced to maintain a relatively tight size distribution across the country. As a working example guideline upper thresholds may be set at twice the 2001 average size. Where these thresholds are exceeded, OAs and SOAs will be considered for splitting, enabling a simple lookup to be maintained between 2001 and 2011 areas. Complex structures and large institutions will not be split. OAs containing them will remain above these thresholds. Research to date suggests that < 5% of OAs and < 3% of SOAs will need to be considered for splitting.

5c. Whilst maintaining stability, changes may be allowed to fix locally identified existing problems with the OA / SOA hierarchy. Such a possibility will depend on the views expressed in this consultation. Criteria would be established for assessing requests for changes to be made to the hierarchy. Decisions would be largely based on how such changes would affect the data. This review and consultation is the main vehicle for making such requests. A further opportunity may be given prior to the next Census if there is likely to be scope for further changes to be made. Where criteria are met, the task of redrawing will be owned by ONS.

[See section 6.5]

6. OAs and SOAs likely to require maintenance will be investigated at the time of the Census. Splits, mergers and any boundary amendments will be planned. Planned amendments will be confirmed with actual Census 2011 data when any final changes will be made. This will cause no delay to outputs and will be ahead of the release of any small area data. There will be no maintenance prior to the next Census in the absence of definitive data to base it on.
7. Lookups will be produced to identify all changes to OAs and SOAs following maintenance.
8. OAs will not be amended to reflect changes to the underlying postcode geography. Following maintenance changes, a postcode to OA lookup will be made available, flagged as 'best fitted' where necessary.
[See section 6.6]
9. Communal establishments will not be split out into their own Output Areas.
[See section 6.7]
10. Work will be undertaken to assess the potential for aligning synthetic OA boundaries to topographic features. This work will have a prime concern that there should be no detrimental impact to digital boundary

licensing terms. It is ONS' strongly held line that OAs and SOAs will remain freely accessible.

[See section 6.9]

11. An Upper Layer of SOAs will not be created unless there is sufficient benefit identified for doing so through this consultation.

[See section 6.11]

12. Following the 2011 Census, OAs will be recoded to the same coding scheme used for SOAs. Lower and Middle Layer SOA codes and labels will remain unchanged, unless areas are impacted by maintenance change. Local alternative names will be supported on lookups where requested.

[See section 6.13]

4 Small Area Geography Policy Consultation Questionnaire

A. About You

1. What is your name? _____

2. Which of the following best describes the organisation that you represent?
[please tick one box only]

No organisation (member of the public)

→ Go to question 4

Central Government

Local Government & Partner Organisations

Government Statistical Agency

Neighbourhood Renewal

Academia

Commercial Sector

Community Group

Health Sector

Other

→ Please specify:

3. What is the name of your organisation?

4. Are you willing for ONS to contact you, to explore your answers further?

Yes No

If yes: Telephone _____

e-mail _____

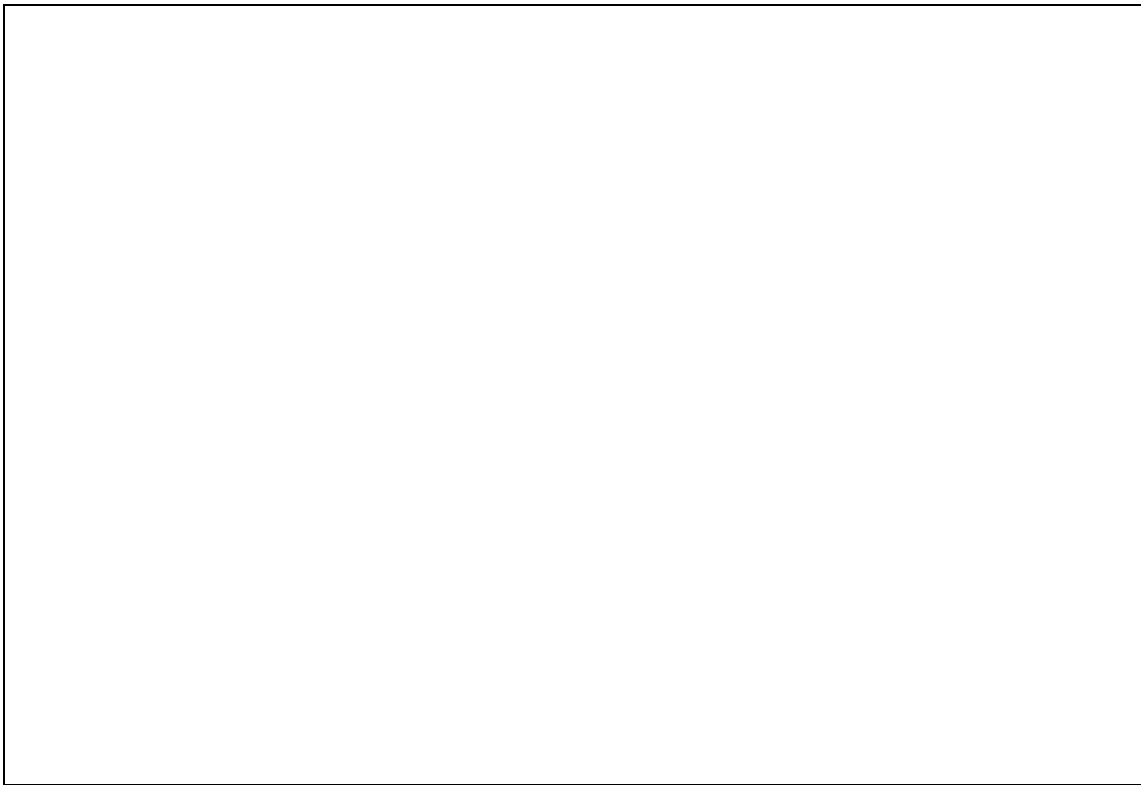
Address _____

Please note, all responses to the consultation will be made public.

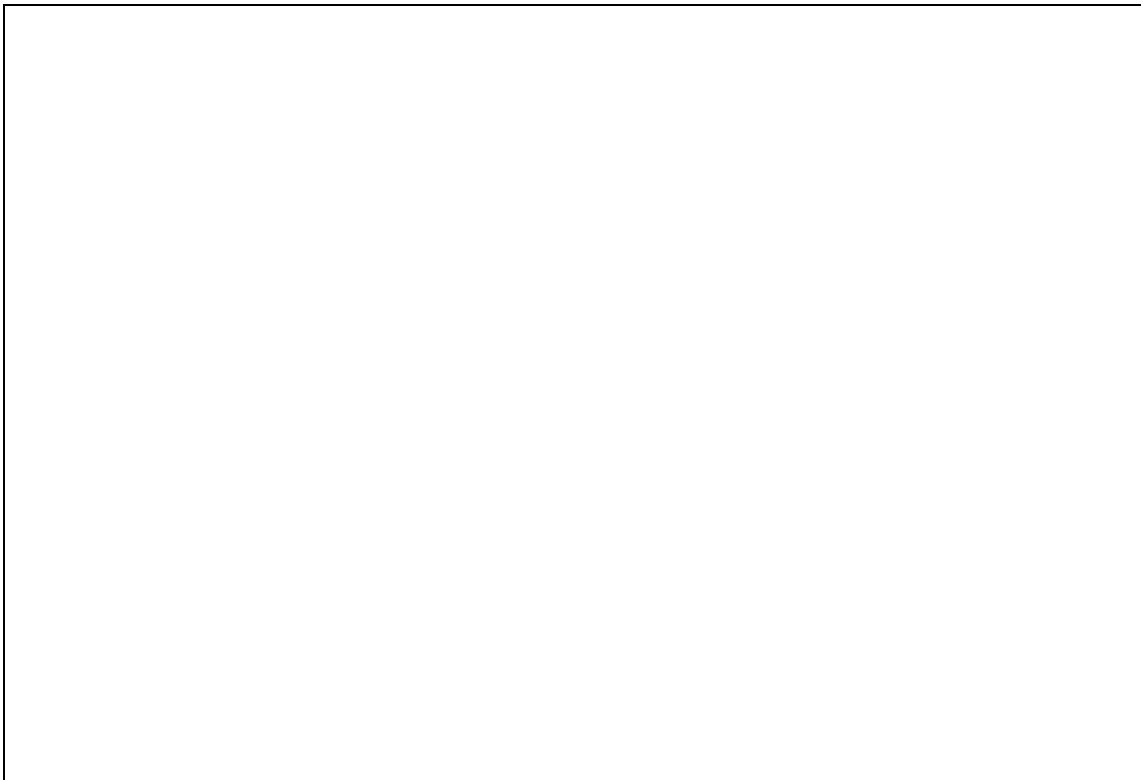
When answering questions, please continue writing on a separate sheet where necessary.

B. Reviewing Output Areas and Super Output Areas

5. For what purposes have you used OAs and SOAs?



6. How have you used OAs and SOAs?



7. How useful have you found OAs and SOAs in terms of:

- a. analysing census data?
- b. analysing other datasets?
- c. using the Neighbourhood Statistics service?
- d. building other geographies of interest?
- e. matching different datasets?
- f. taking action as a result of data analysis?

8. Can you see any further potential uses for a stable small area geography?

Yes No

If yes,

- a. what potential uses can you see?
- a. are there any technical issues preventing this potential use?
- b. are there any other issues preventing this potential use?

9. Are there any OAs or SOAs that you have found to be particularly problematic during data analysis?

Yes No

If yes,

- a. which areas are particularly problematic?
- b. why are they problematic?

10. Would it be useful to you if small area data was also released for geographies other than OAs and SOAs?

Yes No

If yes;

- a. which one other geography is of most interest to you?
- b. which other geographies are also of interest?
- c. which datasets are of most interest?
- d. what key benefits would this bring?

C. Policy and Design

For questions 11 - 25, please indicate your level of agreement with the statement given by circling a number from 1 to 5. On this scale:

- 1 = Disagree strongly
- 2 = Disagree moderately
- 3 = Neither agree nor disagree
- 4 = Agree moderately
- 5 = Agree strongly

11. For statistical purposes, a stable small area geography is important.
[see sections 5 and 6.1]

Disagree strongly – 1 2 3 4 5 – Agree strongly

Please note any further comments.

12. The OA / SOA hierarchy is a useful small area geography.
[see sections 5 and 6.1]

Disagree strongly – 1 2 3 4 5 – Agree strongly

Please note any further comments.

13. It is important to have a nationally consistent small area geography across England and Wales.

Disagree strongly – 1 2 3 4 5 – Agree strongly

Please note any further comments. If applicable, please explain which aspects of consistency are important.

14. It is important to have a nationally consistent small area geography across the whole of the UK (England, Wales, Scotland and Northern Ireland).

Disagree strongly – 1 2 3 4 5 – Agree strongly

Please note any further comments. If applicable, please explain which aspects of consistency are important.

15. It is important to have a small area geography that is consistent between Census 2001 and Census 2011.
[see section 6.1]

Disagree strongly – 1 2 3 4 5 – Agree strongly

Please note any further comments.

16. The OA and SOA hierarchy should be completely redrawn from scratch.
[see section 6.1]

Disagree strongly – 1 2 3 4 5 – Agree strongly

Please note any further comments.

17. If the OA and SOA hierarchy were completely redrawn from scratch, it should be based on:
[see sections 5 and 6.1]

a. whole postcodes that exist at the time

Disagree strongly – 1 2 3 4 5 – Agree strongly

b. 'hard' physical features (e.g. roads, rivers, railways)

Disagree strongly – 1 2 3 4 5 – Agree strongly

c. administrative boundaries that exist at the time

Disagree strongly – 1 2 3 4 5 – Agree strongly

d. neighbourhood definitions, where they exist

Disagree strongly – 1 2 3 4 5 – Agree strongly

e. grid squares

Disagree strongly – 1 2 3 4 5 – Agree strongly

f. Please note any further comments

18. Small area geographies for statistical purposes should have tight population ranges.
[see section 6.2]

Disagree strongly – 1 2 3 4 5 – Agree strongly

Please note reasons and any further comments.

19. Small area geographies for statistical purposes should have tight household size ranges.
[see section 6.2]

Disagree strongly – 1 2 3 4 5 – Agree strongly

Please note reasons and any further comments.

20. Large communal establishments should be separated from surrounding areas in a small area geography.
[see section 6.7]

Disagree strongly – 1 2 3 4 5 – Agree strongly

Please note reasons and any further comments including, if applicable, how a 'large communal establishment' should be defined.

21. The design of small area geographies should allow boundaries to be drawn around unpopulated land.
[see section 6.10]

Disagree strongly – 1 2 3 4 5 – Agree strongly

Please note reasons and any further comments.

22. There would be value in establishing an Upper Layer of SOAs.
[see section 6.11]

Disagree strongly – 1 2 3 4 5 – Agree strongly

Please note any further comments including, if applicable, details of the benefits that an Upper Layer would bring.

23. There would be value in establishing a separate small area geography for reporting business data.
[see section 6.12]

Disagree strongly – 1 2 3 4 5 – Agree strongly

Please note any further comments including, if applicable, details of the benefits that such a geography would bring.

24. The OA and SOA boundaries that currently exist should now be neatened, where possible, to underlying topographic features (e.g. by snapping them to building and open land boundaries, roads, rivers etc).
[see section 6.9]

Disagree strongly – 1 2 3 4 5 – Agree strongly

Please note any further comments including, if applicable, which features or datasets (e.g. OS MasterMap) to use.

25. The current licensing regime for OA based digital boundaries is acceptable.
[see section 6.9]

Disagree strongly – 1 2 3 4 5 – Agree strongly

Please note any further comments

For questions 26 and 27, please indicate your answer by circling a number from 1 to 5.

26. How important is it for you to have freely available digital boundaries for small area statistical geographies?
[see section 6.9]

Not important at all – 1 2 3 4 5 – Very important

Please note any further comments

27. How important is it for SOAs to be given more meaningful names?
[see section 6.13]

Not important at all – 1 2 3 4 5 – Very important

Please note any further comments

28. Please comment on any other topics that you feel are relevant.
[on a separate sheet if necessary]

APPENDICES

5 Background

5.1 2001 Census – Output Areas

Output Areas (OAs) were created for 2001 Census outputs. Previously, the geography of census data collection, Enumeration Districts, had also been used as the geography of small area outputs for England, Wales and Northern Ireland. Scotland had already developed a set of OAs for 1991 outputs.

Users of census data had various requirements for small area outputs from the census. The main requirements were as follows:

- a geography aggregating to exact local government units (e.g. wards & local authorities);
- a geography based on postcodes / postal geography (frequently the only geo-referencing field in many other databases and datasets);
- a geography that matched that of the previous census;
- a geography that could be related to environmental data (e.g. grid squares);
- a geography that would be freely available.

In many ways these requirements were competing. The OA solution satisfied the first two and the last of these requirements (for non commercial use) and was enabled by GIS developments together with the grid referencing of 2001 Census records. Thiessen polygons were drawn around Census addresses. These were then aggregated to postcodes, with internal boundaries dissolved, to form unit postcode polygons. An iterative zoning algorithm was applied to group these postcode polygons together into larger zones (the OAs), nested within ward and parish boundaries. Where postcodes straddled ward and parish boundaries, they were split between OAs. Within these geographical parameters the zoning algorithm was designed to create OAs above a minimum population and household threshold size (100 & 40 respectively), tightly grouped around a target household size (125). It also tried to create OAs with a compact shape and with a degree of homogeneity in terms of housing tenure and type.

Within these parameters and targets, again sometimes competing, the solution was optimised as far as possible. It was broadly seen as a success, although some criticisms concerning the resulting abstract nature of OA boundaries were raised.

The potential to release census data for other small area geographies was restricted by issues of data confidentiality. Whilst minimum population thresholds can be set for individual geographies, the release of data for an overlapping geography gives rise to the possibility of subtracting one set of tables from the other to reveal statistics for sub-threshold areas. This is known as disclosure by differencing and is primarily a problem where an area or areas of geography A actually nest within an area or areas of geography B. Very small area geographies increase the risks of disclosure by differencing.

At the small area level, Census 2001 data was released for OAs and for wards that existed in 2003 (as OAs nested within those wards). The potential also existed to release data for other aggregations of OAs.

5.2 Neighbourhood Statistics – Super Output Areas

Neighbourhood Statistics were created within ONS as part of the National Strategy for Neighbourhood Renewal. From a geographical perspective, the report of Policy Action Team 18: “Better Information” (PAT 18, published 2000) stated:

To make neighbourhood renewal work in small areas requires data to be available broken down by ward at the very least The organisational boundaries of data are varied, and do not match each other Area-based programmes are often focussed on smaller areas that may not match either these or ward boundaries Ward boundaries change substantially from year to year, making geographical comparisons over time difficult.

If data could be referenced at the smallest geographical level (such as a grid reference) it then could be aggregated to higher geographic levels and, in principle, to any pre-defined boundary. This would allow greater flexibility in compiling local area information. It would also open up the possibility of realistic comparisons over time, by allowing data to be recast to new geographic areas as boundaries change.

Such a reliable geographic referencing framework is missing in this country insofar as there is some positive activity, the task of building it is fragmented between agencies. The most comprehensive geographic tool available at the moment is the postcode. Postal geography is useful, but not ideal, for referencing residential locations, is weaker when locating businesses and business activity and is particularly weak when locating activities conducted by the voluntary sector.

However, while there are considerable benefits in using geographic referencing as a basis for compiling area information, some limits will be needed to prevent figures being produced that are so small that they could disclose information about an individual, therefore breaching confidentiality.

The first release of the Neighbourhood Statistics service was based on ward level data aggregated to frozen 1998 ward boundaries. However, as well as being subject to substantial change from year to year; wards also vary hugely in size across the country. Indeed, in terms of population, wards range from under 1,000 to well over 30,000. Therefore, not only are comparisons over time difficult (without fixing on historic boundaries), but comparisons over space are also difficult. In areas such as Birmingham, with average ward sizes well in excess of 30,000 people, the granularity of the geography would be difficult to label as “small area” (see Fig. 1).

In early 2003 ONS proposed a new layered geography for Neighbourhood Statistics. The smallest zones would be OAs and the largest zones would be local authorities. Between these, intermediate layers built from aggregations of the layer below would be constructed. These intermediate layers were termed Super Output Areas.

Feedback to the proposals was broadly positive. In 2004 both a Lower Layer and Middle Layer of SOAs (LSOAs & MSOAs) were published. An Upper Layer was also proposed but has not yet been constructed.

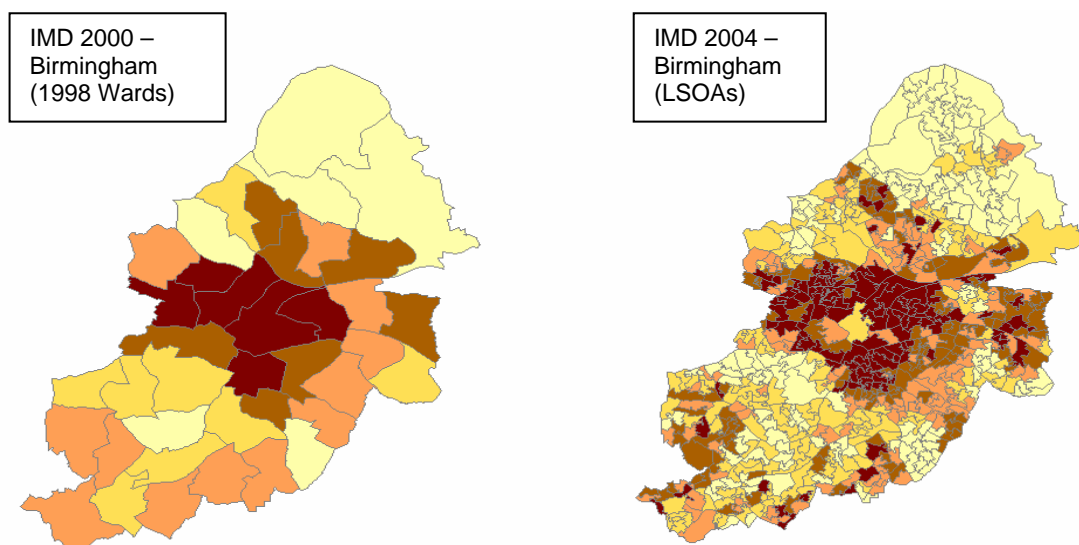
SOAs are the core geography for both the collection and output of Neighbourhood Statistics. It was stated that they would be durable, to provide a consistent geography for readily measuring change over time. As with OAs, they also have relatively tight population ranges within each layer. SOAs are, subject to agreement and confidentiality controls, used for the aggregation of source records by data suppliers and for the release of data on Neighbourhood Statistics. It was proposed that outputs for geographies other than SOAs would be enabled through estimation methods. Using OAs as building blocks gives the potential to allow the ad-hoc aggregation of units to provide estimates for any area with the possibility of disclosure already removed. Data could thus be best fitted or synthetically estimated to other non co-terminous geographies.

As SOAs are aggregations of OAs, and LSOAs nested within Census Standard Table Wards, Census data which had already been released for OAs and ST Wards, could also be released for SOAs. This not only provided the richness of census data for the new geography, but critically it also provided population data for SOAs which could be used as denominators for rates and indicators.

5.3 Implementing PAT 18 Report Recommendations and the OA / SOA hierarchy

The introduction of SOAs has enabled data to be aggregated to this stable and nationally consistent geography. Outputs for SOA data have now overtaken those for other geographies on Neighbourhood Statistics. An early example of the benefits that can be realised from using this geography was apparent with the release of the Indices of Deprivation in 2004. In 2000, this analysis had been released for the approx 8,000 wards in England. Over time these wards have changed, and the size of wards, whilst suitable for identifying pockets of deprivation in some local authorities, were totally unsuitable in others. This is best exemplified by comparing mapped outputs for Birmingham (see Fig. 1).

Fig. 1 – Index of Multiple Deprivation for Birmingham (2000 and 2004)



© Crown copyright. All rights reserved (ONS GD272183 2006).

Note - these maps are colour coded into 5 breaks. The darker the colour, the more deprived the area.

A number of barriers to realising the vision of the PAT 18 report still exist though. They are as follows:

- Legal barriers, in tandem with responsibility for data confidentiality remaining with data suppliers, have prevented ONS from obtaining unit level administrative datasets;
- As a result, disclosure controls and aggregation to LSOAs, or more frequently MSOAs, takes place before data is passed to ONS;
- Whilst OAs can be used as building blocks on Neighbourhood Statistics to create best fitted areas of interest, due to a lack of OA level data there is little other than census data available for these user defined OA aggregations;
- Where data is provided pre-aggregated to LSOAs or MSOAs, it is currently not possible to produce estimates for other geographies to a publishable level of quality;
- In the absence of unit level datasets, the impetus for developing suitable disclosure control methods for enabling the originally envisaged building block approach is missing;
- Very little data from outside central government has been made available to Neighbourhood Statistics;
- The absence of a freely available national spatial addressing infrastructure, means that the majority of non census datasets still rely on the postcode for geo-referencing purposes, and are likely to continue to do so for the foreseeable future.

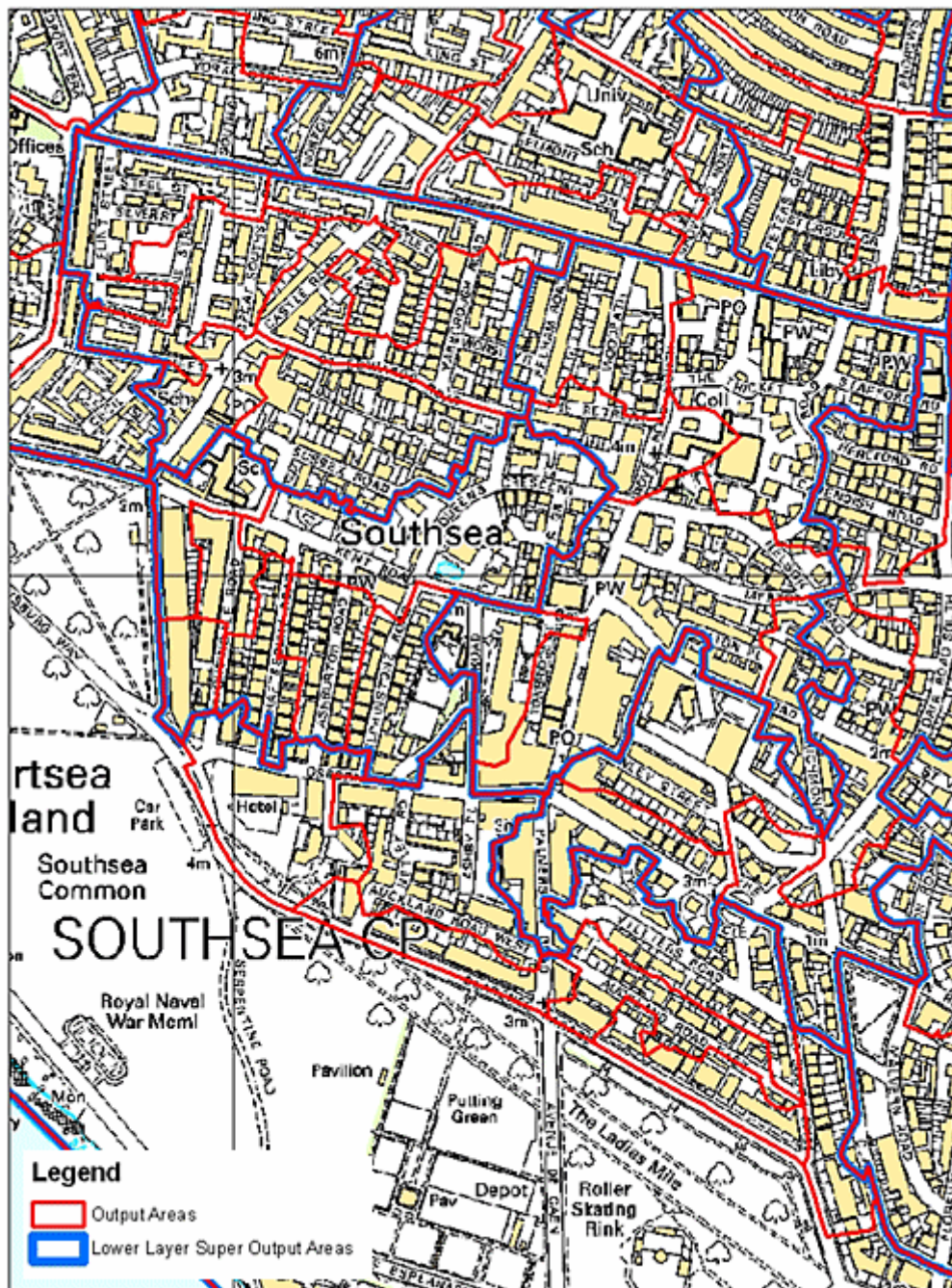
Given these barriers to moving towards the vision of flexibly creating outputs for different geographies, the focus on the principal geography for data aggregation becomes ever greater for local analysis purposes. The benefits of having a geography that is consistent across the country, and that is

intended to be stable through time, are widely recognised. The rate of adoption of SOAs has slowly increased in the user community, but there are understandable criticisms.

An acknowledged criticism of OAs and SOAs has been recognised since they were first introduced: OAs were built from artificial postcode polygons, which are an abstract 'soft' geography. At the time there were no digital postcode polygons and even those that do exist now have been created by a similar process to that used in the creation of OAs. As a result, the boundaries do not relate well to 'hard' geographical features, such as transport links or other elements of physical geography. They were designed to ensure tight population ranges, whilst keeping areas of housing homogeneity together as far as possible. They do not necessarily delineate areas that users might think of as being discrete communities or areas of interest, and were never intended to. This aspect of their design becomes more of an apparent drawback when taken in tandem with an inability to produce outputs for other geographies of interest.

Figures 2 and 3 give examples of OAs and SOAs related to the underlying geography. Figure 2 shows OAs and LSOAs in an urban area, Southsea. Figure 3 shows OAs, LSOAs and MSOAs in a more rural area, Anglesey.

Fig 2. – OAs and LSOAs in Southsea



© Crown copyright. All rights reserved (ONS GD272183 2006).

Fig. 3 – OAs and SOAs in Anglesey



© Crown copyright. All rights reserved (ONS GD272183 2006).

6 Topic notes

6.1 Stability or re-drawing

A fundamental issue in determining the direction of the policy is whether we should aim for stability in output geographies for small area data analysis purposes. When OAs were designed they were not intended to form the basis of a long lived stable geography. But SOAs were designed with this intention, and were built from groups of OAs. Since their introduction, feedback on this issue has been mixed.

The main point behind the introduction and intended use of SOAs is that they will provide a consistent basis for collecting and disseminating information. They are designed for this purpose and intended to eliminate the problems brought about by the previous use of administrative boundaries. A policy of stability allows data to be tracked and compared through a significant period of time on a consistent geographical basis.

Whilst OAs were not originally designed with longevity in mind, the potential to use them again for Census 2011 outputs brings a real opportunity: geographical consistency between censuses at a small area level for the first time. OAs are the basis of SOAs, and so a move away from them would in turn undermine the SOA hierarchy. However, there are advocates of the same postcode based algorithm being run again following the next Census to create a new set of OAs.

It is accepted that OAs and SOAs will often have to be a best fit building brick geography. Due to their small size, their suitability as building bricks should continue to hold in the future, much as postcodes have always been used. The needs of those requiring an evidence based approach are likely to be better met by consistency in small area building bricks, rather than by revising the geography every few years. Over time, longevity should bring familiarity, greater use, and an incentive for investment in data management processes.

The point has been strongly put forward that areas could be designed that made more sense on the ground, and that were recognisable. However, within the design parameters set, this would be a significant task. There is no unique definition of local areas and different agencies would have different ideas for creating better boundaries. The question of whether boundaries are recognisable is also highly subjective.

Alternative approaches to the design of statistical geographies are evident in other countries. One such approach is the 'block' design. A 'block' is an area of land entirely surrounded by streets, streets and coastline, or other 'hard' edges such as railway lines, rivers or canals. They need only change when significant physical change has occurred in the environment. For these reasons a block design is a very attractive alternative, not only for data reporting but also for enumeration purposes and the analysis of other geographical datasets, such as address lists. It may also be more suitable for adoption by others constructing local boundaries for administrative, policy or

service delivery reasons, potentially enabling a degree of geographical compatibility over time.

The attractions of a block based design are tempered by a number of hurdles, many specific to the current situation in the UK. We would propose that all of these hurdles, outlined below, need to be negotiated in order to consider such a redrawing and the complete data break that would result.

Blocks would vary more in size than OAs, resulting in a broadening of resident population and household size ranges. In order to be useful for statistical purposes, they would need to be split or amalgamated to form more practical reporting units. This would force some difficult choices to determine which blocks to split or group together, and the criteria for doing so.

One potential criteria for splitting and amalgamating blocks would be based on postcode boundaries. But this would compromise the design, leaving it still tied to an abstract geography built and regularly maintained for the sole purpose of delivering mail. The driver for trying to build in a postcode respecting design, is to allow address level and postcode level data to be aggregated to it. Whilst postcodes are a conveniently available geography, they are not the firm foundation we would ideally want to base statistical data and statistical geographies on. The need for a link via postcodes from unit level data to data aggregation unit is embedded in UK data, but ideally it must be broken. This will only be possible with the advent of a widely available and easily accessible national addressing infrastructure. Such a development would enable a direct lookup link to be established from addresses or buildings to a statistical geography, and is an essential requirement.

Digital block boundaries would need to be derived directly from digital topographic data. As a result there would be significant data licensing issues to negotiate in order to make these boundaries widely accessible, and ideally freely available.

Blocks would need to be aggregated to create a layered hierarchy for the release of data to a standard statistical geography. As with OAs and SOAs, blocks and aggregations of them would need to be spatially constrained by other geographies. These higher geographies themselves should be stable enough to form part of the standard statistical geography. OAs and LSOAs were constrained by wards that existed at the time. But, for the reasons already outlined, the use of wards as a basis for a statistical geography constrains rather than enables data analysis. The likelihood of getting a broad consensus on what geographies, or concepts of communities, should be used as the spatial parameters for an alternative geography will be small. Again, however, it is essential to achieve.

Views on the desirability, or otherwise, of maintaining a stable OA / SOA geography are sought. The design requirements for a small area statistical geography remain as diverse as those identified in section 5.1. No single geography can meet all of these competing requirements. This is especially so given current and foreseeable UK geographical data frameworks. Whilst this is the case, the concept of SOA stability provides an existing anchor point for data analysis across time and space. SOAs deliver the requirements of a

statistical geography within the data limitations that currently exist. OAs also give the potential for a wealth of data to be best fitted to other geographies of interest, providing the hurdles of making data available to ONS at this level can be negotiated.

6.2 Scale of real world maintenance

Proceeding with a policy of stability introduces a requirement to maintain the geography. This is because of a need to respect minimum size thresholds and a desire to maintain a reasonably constrained size range within each layer of the hierarchy. When designing the geography, minimum thresholds were set for both population and household counts for each layer in the hierarchy. Output Areas were not actually constrained by maximum thresholds when they were designed, rather the algorithm that optimised their design did so by minimising their deviation from a target size.

Stability clearly leads to a desire to minimise the scale of maintenance that will be required. We would aim to keep the proportion of areas where maintenance was required to under 5% of the total. The question then becomes one of assessing the scale of change that is likely to be required for what thresholds. But, in the absence of a census itself, we currently only have small area population estimates up to 2004 to use. These do not go down to OA level.

To try and clear this hurdle we have devised a proxy household count by analysing delivery points by postcode in the National Statistics Postcode Directory (NSPD). These have been calibrated against census results and cross referenced against council tax registers. We have then rolled the counts forward year by year to 2006 with the results shown in the tables below.

Table 1 - Output Areas (England and Wales)

Year	OAs with zero dwellings	OAs with less than 30 dwellings but more than zero	OAs with 30 - 40 dwellings	OAs with more than 250 dwellings	OAs with more than 300 dwellings
2001	29 (0.017%)	92 (0.052%)	279 (0.159%)	275 (0.16%)	109 (0.062%)
2002	364 (0.207%)	181 (0.112%)	290 (0.277%)	319 (0.18%)	128 (0.073%)
2003	29 (0.017%)	93 (0.053%)	253 (0.197%)	518 (0.30%)	218 (0.124%)
2004	32 (0.018%)	81 (0.046%)	241 (0.184%)	830 (0.47%)	387 (0.221%)
2005	31 (0.018%)	91 (0.052%)	237 (0.187%)	1107 (0.63%)	566 (0.323%)
2006	37 (0.021%)	93 (0.053%)	185 (0.158%)	1933 (1.1%)	823 (0.469%)

Note - average OA resident household size (2001) = 125

The figures in Table 1 show that the number of Output Areas falling below threshold or with zero population is very small indeed, and that the majority of such cases are actually very close to the threshold value. The number of

such areas seems to be quite stable. Noise in the NSPD data such as lags in address count updating, geo-referencing revisions and overlap of terminated and newly introduced postcodes also means that robustly identifying below threshold areas is difficult.

Further investigation revealed that many of these examples are caused by temporary postcode changes or coordinate referencing problems. For example, the temporary increase in zero dwelling count figures in 2002 is almost certainly a result of the large-scale postcode revision that took place in that year rather than a reflection of any real change on the ground. The stability of the other annual figures suggests that the true number of zero count cases is closer to 30 Output Areas. These tend to persist through time, and investigation of some of them suggests that it is likely that they are a result of geo-referencing imprecision in Output Area boundaries, mainly caused by the use of postcode unit Thiessen polygons as a basis for Output Area generation, rather than a reflection of actual zero dwelling counts.

Of more interest for maintenance policy formulation is the relatively larger group of Output Areas that have more than twice the original target number of dwellings (250) and are therefore potential candidates for splitting into smaller units. The number of such areas increases through time, which we would expect given a rising population and the shift towards smaller household sizes in England and Wales. By 2006, about 1.1 per cent of Output Areas (almost 2000 areas) fall into this category. If it is assumed that the changes that we see in the data for 2001-2006 continue in a similar way from 2007-2011 then approximately 5 per cent of areas may contain more than 250 dwellings by 2011. However, there is no guarantee that this trend will continue: the number of time points available is really too small to put forward a strong argument supporting this assumption, and the use of dwelling related data is not a perfect proxy for persons or households.

Ground truth in known areas of development suggests that our method picks up new development well.

Table 2 – Lower Layer Super Output Areas (England and Wales)

Year	LSOAs with zero dwellings	LSOAs with less than 400 dwellings but more than zero	LSOAs with more than 800 dwellings	LSOAs with more than 1200 dwellings
2001	0	118 (0.343%)	3,556 (10.344%)	49 (0.143%)
2002	22	197 (0.573%)	3,639 (10.585%)	53 (0.154%)
2003	0	103 (0.300%)	3,965 (11.534%)	74 (0.215%)
2004	0	94 (0.273%)	4,329 (12.592%)	112 (0.326%)
2005	0	96 (0.279%)	4,672 (13.590%)	170 (0.495%)
2006	0	86 (0.250%)	5,135 (14.937%)	239 (0.695%)

Note - average LSOA resident household size (2001) = 600

The number of LSOA areas affected by threshold problems is small. Areas below the minimum design threshold of 400 decrease in number through time, and in 2006 only 0.25 per cent of areas are candidates for maintenance.

Assuming this trend continues, the number of candidates will have declined further by 2011.

The number of areas that contain approximately double the average number of households (1200) at LSOA level increases through time, and by 2006 has reached 0.695 per cent of the total. If this trend were to continue then by 2011 the number of LSOA areas that might require maintenance would be of the order of 2.5 per cent.

When designing LSOAs a permeable maximum threshold of 800 households was set. But a significant minority of LSOAs exceeded this threshold as they could not be split further into sensible subdivisions all above threshold.

Table 3 – Middle Layer Super Output Areas (England and Wales)

Year	MSOAs with zero dwellings	MSOAs with less than 2000 dwellings but more than zero	MSOAs with more than 4000 dwellings	MSOAs with more than 6000 dwellings
2001	0	29 (0.366%)	703 (8.883%)	7 (0.088%)
2002	0	50 (0.632%)	708 (8.946%)	6 (0.076%)
2003	0	18 (0.227%)	766 (9.679%)	6 (0.076%)
2004	0	14 (0.177%)	814 (10.286%)	9 (0.114%)
2005	0	11 (0.139%)	865 (10.930%)	11 (0.139%)
2006	0	7 (0.088%)	953 (12.042%)	16 (0.202%)

Note - average MSOA resident household size (2001) = 3000

As with LSOAs, the number of MSOA areas affected by threshold problems is also small. Areas below the minimum design threshold of 2000 decrease in number through time, and in 2006 only 0.088 per cent of areas are candidates for maintenance.

The number of areas that contain approximately double the average number of households (6000) at MSOA level are equally small, and by 2006 has only reached 0.2 per cent of the total. If this trend were to continue then by 2011 the number of MSOA areas that might require maintenance would be under 1 per cent.

When designing MSOAs a permeable maximum threshold of 4000 households was set. But, as with LSOAs, a significant minority of MSOAs exceeded this threshold.

6.3 Method of applying maintenance for real world changes

We propose applying a method of simple splits and mergers for those OAs and SOAs requiring maintenance. We would use proxy data at the time of the census to plan those areas likely to require maintenance. Once initial census counts are available within ONS, any necessary amendments to planned maintenance changes would be made. Maintenance would not take place before the next census.

Preferred methods to be used are currently under investigation. MSOA splits are likely to be along existing LSOA boundaries, with LSOA splits along existing OA boundaries. Options are still being considered for OA splits. Mergers will be with neighbouring areas within the boundaries of the parent layer. In other words an OA would be merged with a neighbouring OA within the same LSOA.

It is proposed that such maintenance be carried out within ONS in order to eliminate any chance of a subsequent delay to census outputs.

6.4 Alignment to Local Authority boundaries

The OA – SOA hierarchy nests within Local Authority (LA) boundaries. It is proposed that any LA boundary changes implemented in the intercensal period be respected during the Census maintenance process. In such an event, the boundaries of any affected OAs and SOAs would be redrawn to new LA boundaries. Where necessary, OA and SOA fragments that resulted would be merged with neighbouring areas within the new LA boundaries.

This proposal is yet to be tested within ONS.

6.5 Structural maintenance

It is proposed that within the overall limit of no more than 5% of all areas being amended for maintenance purposes, Local Authorities will be able to propose areas that they would like to see redrawn, or where they would like to see the hierarchy amended, because of clear problems in current utilisation. Requests for redraws should not be made simply because of local knowledge of where population has rapidly grown or shrunk, as these areas will be picked up by the normal maintenance process.

Criteria will be established for assessing any requests made to ONS for such redraws. The number that can be accommodated will be determined by the scale of maintenance required for real world change.

It should be noted that any areas structurally redrawn as a result of successful requests, will lose their consistency through time unless such redraws only consist of simple splits and mergers of existing OAs and SOAs. This should be a key consideration before making requests.

6.6 Links to postcodes

OAs were originally designed as being comprised of whole unit postcodes, except where postcodes were split at ward and parish boundaries. Postcode re-organisations are frequently undertaken by Royal Mail. Approximately 1 – 2% of all postcodes are either introduced or terminated each year.

Taken as a whole, the 2011 postcode geography will still match that of 2001 relatively well. Not all postcode introductions and terminations actually result in a shift in physical location, some are just recodes. However, significant changes will have taken place, often localised.

It is recognised that postal geography can be useful for statistical purposes as it is still frequently the only geo-referenceable field in many datasets. However, it is by no means ideal, and ONS policy is to encourage a move to lower level geo-referencing using addresses and grid references.

We propose that no attempt is made to re-align OA boundaries to re-organised postal geographies. Instead the degradation in the link over time that is already taking place, will be allowed to carry on unchecked.

6.7 Communal establishments

OAs currently incorporate all communal establishments in England and Wales. That is to say that even OAs containing very large communal establishments also include other surrounding properties. Some feedback was received, both during the OA creation process and subsequently, requesting that larger communal establishments be isolated in their own OAs.

We are not currently proposing to do this, but would welcome views on whether this is still seen as a valuable development. If so, what would the definition of a 'larger communal establishment' be? One criteria would have to be that the establishment were above the minimum OA population threshold.

6.8 SDC methods and overlapping geographies

As described in section 5.1, disclosure by differencing is a significant barrier to producing actual count outputs for alternative overlapping geographies. The result of this is that a policy of small area geographical stability limits the ability to produce actual count outputs for other non coterminous geographies (for example 2011 Wards).

At this stage in the census planning cycle, however, it is possible to investigate potential disclosure control methods that may enable outputs to be provided for an alternative small area geography.

Feedback is requested on this issue, in particular whether there is a demand for specific outputs on an alternative geography. If there is a demand for another geography, what geography is that?

In terms of the production of non census data for overlapping geographies, again this possibility theoretically exists. However, very few historic datasets are geocoded to address or grid reference level. Even where they are, such geocoding is not necessarily accurate or consistent. Legal barriers to data sharing also prevent ONS being given access to suitably low level referenced data.

6.9 Boundaries and Licensing

OA and SOA boundaries follow artificial thiesen polygons around addresses. As such, they are abstract and do not follow physical features. This means that they frequently cut across non census addresses, transport links, and areas of open ground, where in some instances they need not.

OAs and LSOAs were constrained to Census Standard Table Ward boundaries. As a result they were partially derived from Ordnance Survey data. To enable OA based digital boundaries to be freely available to end users, ONS paid OS a one off licensing fee to cover royalty costs for 10 years. ONS has provided free OA and SOA boundary digital boundary products since the geographies were introduced. However, their use does require some GIS expertise.

The licensing fee ONS paid OS did not permit the commercial use of OA and SOA boundaries. Indeed the commercial sector has so far failed to reach an agreement with OS on this issue. This has resulted in a situation where Census distributors provide Census data but cannot provide the geographical boundaries that the data are aggregated to. Similarly, OA and SOA boundaries have not been incorporated into any third party packaged products or client solutions. End users still have to come to us, and still need GIS expertise to use our boundaries.

To better relate OA boundaries to the real world, we are keen to investigate the potential for realigning boundaries to an underlying ground truth topographical layer, such as OS MasterMap. This would result in small changes to the boundaries that would not affect the stability of statistics for areas.

Such a realignment would mean the boundaries were more clearly derived from OS data. It would also potentially enable OS to include OA and SOA boundaries as a data layer in MasterMap. We would not, however, proceed with boundary realignment if it were likely to be detrimental to our current or future licensing terms, or our ability to freely disseminate boundaries to third parties. OAs and SOAs are intended to be for the national good, as a geography to enable consistent small area statistical analysis.

We have carried out initial investigations into the realignment of OA boundaries to topographic features. These have shown mixed results. Whilst boundaries can be realigned around features that they currently bisect, the automation of such a process is by no means straightforward. It is still frequently necessary to break topographic boundaries to maintain the integrity and contiguity of OAs. Visually also, realigned boundaries can appear more 'jagged' than they were before as what were previously straight, albeit artificial lines, become much more complex when snapped to and around the features they intersect on the ground. Further investigations in this area have been halted until there is more clarity on the issue of licensing. At present we feel it is unlikely that OAs and SOAs will be realigned before Census 2011.

A redrawn OA geography based on physical features, such as the 'block' geography approach, may eradicate some of the technical and visual issues noted above. However, the issue of licensing would be brought even more sharply into focus, as the design would be based on topographic features.

We are keen to seek user views on: the desirability of realigning OA boundaries to an underlying topography; the current licensing regime; and any problems users have had in using OA and SOA digital boundary files.

6.10 OA / SOA contiguity and zero population OAs

Both OAs and SOAs were designed to be contiguous and space filling. That is to say that each area forms one whole polygon only, and the total land surface is contained within. An area cannot be made of two or more distinct parts with another area or areas in between.

Statistical geographies in some other countries have introduced the concept of areas with zero population within them. In less densely populated countries than England and Wales, especially those with tracts of wilderness, this is an understandable way of ensuring that the whole territory is covered to enable the reporting of non population based statistics.

The concept of zero population OAs may, however, be considered to have some merit here too. There are rural areas where the boundaries of OAs can extend for some distance simply to include the whole land area, whilst there is no resident population actually present or only a very sparse resident population. There are also instances in urban and urban / rural fringe areas where the same is true for large industrial parks.

We would like to explore user views on this topic. There are significant issues to consider, including OA contiguity and the possibility of introducing differencing risks as a result. An alternative would be to publish a separate 'unpopulated' layer that could be overlaid on top of OA boundaries. Such a layer might be based on population sparseness. Would such a development be seen as being of use?

6.11 Upper Layer SOAs

The intention has always been to create an Upper Layer of SOAs in consultation with key stakeholders. It was envisaged that an Upper Layer would have minimum size thresholds of 10,000 households and 25,000 people. USOAs would be aggregations of MSOAs, and as such may enable the release of datasets that currently cannot be released below Local Authority level.

Since the Lower and Middle Layers were introduced, little demand has been expressed for an Upper Layer. As a result, we do not plan to create and release one. However, as part of this consultation, users and stakeholders are requested to identify any clear requirements for an Upper Layer that they

may have. Any benefits identified will be assessed and the decision not to create an Upper Layer will be reviewed.

6.12 Business Output Areas

OAs and SOAs are designed primarily for social statistics. They are based on population and household size ranges. They do not take into account economic or business criteria in their design, such as the number of businesses or employees. Whilst they can still be used for economic statistics, this is not their intention.

The majority of economic data is reported at the regional level or higher, and certainly no lower than at local authority level. However, if there were interest in having analysis at a more localised level and datasets that would support it, this may be a reason for considering an Upper Layer of SOAs. In this case factors such as the number of businesses could be investigated in its creation.

At the micro OA level, basic analyses such as the number of businesses or employees present may be useful. However, due to the nature of OA design, areas where there are an extremely high number of businesses and employees present (such as the City of London) are also areas where the resident population is relatively widely dispersed in comparison. The net effect of this is that the majority of OAs have very few businesses within them, whilst a tiny minority have vast numbers of businesses and employees present. It would be theoretically possible to split these social OAs into smaller business OAs.

Views on the desirability of investigating the potential for introducing economic design criteria for either an Upper Layer of SOAs or for sub-dividing OAs are sought.

6.13 Naming and coding

All OAs and SOAs carry a unique code. OA codes include embedded ward intelligence, which is now becoming out of date. SOA codes intentionally contain no other embedded intelligence in order to future proof them.

We propose that OAs will be recoded into the same format as SOA codes following the 2011 Census. A lookup from old to new codes will be provided.

We propose that SOAs that remain unaltered following maintenance will keep their existing codes. Maintained SOAs will be given a new code according to the current coding scheme. Their pre maintenance codes will no longer be used. A lookup of maintained SOAs will be provided.

OAs are not labelled. SOAs are currently labelled based on the Local Authority in which they nest. For example MSOAs in Westminster are labelled Westminster 1, Westminster 2, Westminster 3 etc. LSOAs within Westminster 1 are labelled Westminster 1a, Westminster1b, Westminster 1c etc.

When released, SOAs were intentionally not given more meaningful names. The fact that they were not designed around concepts of localities, taken in tandem with their small size, led us to believe that any attempt to name them would be a significant and potentially emotive task. It could be difficult to reach local agreement. Similarly, if names were tied to place, we anticipated complaints from those not happy with the name given to the SOA in which they lived. But there is a view held by many that the advantages of naming outweigh the potential disadvantages. Where alternative names have been created and used for the local analysis of SOA data, we have been happy to include these on lookup files. We are only aware of a handful of such examples. Alternative names are not used for reporting Neighbourhood Statistics.

7 Contact Details

The on-line questionnaire and blog forum is available at:

www.statistics.gov.uk/geogconsult

Further details can also be found at:

www.statistics.gov.uk/about/consultations

E-mailed responses and questions should be sent to:

info@statistics.gov.uk

Posted responses should be sent to:

Geography Consultation
Customer Contact Centre
Room 1.015
Office for National Statistics
Cardiff Road
Newport
NP10 8XG

You can contact us via our customer contact centre on:

+44 (0)845 601 3034