Institution: University of Southampton

Unit of Assessment: 17B Geography, Archaeology and Environmental Studies

Title of case study: 17B-02 Redrawing the Lines: Developing New Census Output Geographies

1. Summary of the impact

Research by the University of Southampton has led to an entirely new approach to the creation and management of small geographical areas for the publication of official statistics, including those from the 2001 and 2011 UK Censuses and the Neighbourhood Statistics Service. The software at the heart of this transformation is now used in 10 countries, while the academics responsible for it have helped inform government decisions, are integral to the policy and practice of the Office of National Statistics and have presented evidence to various influential committees. The research continues to deliver benefits to a large user community.

2. Underpinning research

University of Southampton research has led to the creation of new geographical areas for the publication of official statistics, including from the Census and the Neighbourhood Statistics Service, so enabling the accurate recording and monitoring of local area data over time.

Before 2001 the Office for National Statistics (ONS) used the same geographic areas both to collect census data and report on findings. These manually-created areas were determined primarily by requirements for data collection, resulting in data being published for areas with diverse population sizes, socio-economic composition and geographical shapes. Together with changes in geographical boundaries between censuses, these widely varying statistical characteristics severely hindered comparative analysis both between areas and through time. In 1991, population counts in approximately 4,000 areas were so small that data could not be published due to confidentiality concerns.

In 1996 David Martin (then Reader in Geography at Southampton; Professor from 2000) began a programme of research hosted by ONS and funded by ESRC **[3. G1]**. Martin developed a conceptual model and prototype geographic information systems (GIS) algorithm for the automated design of a system of small geographical areas, named Output Areas (OAs), for the reporting of Census data, published in **[3.1]** and **[3.2]**.

At the core of this 'automated zone design' approach is an algorithm for the creation of new areas by the iterative recombination of small geographical building-block zones into a number of larger zones. These larger zones are optimised to meet specified statistical and geographical constraints such as minimum population size, social homogeneity and compactness of shape. Rather than the previous approach of using the same areas for data collection and reporting, Martin wrote software to design an optimal set of statistical zones for reporting data once the Census information had been collected.

This work was adopted by the ONS to create OAs for reporting findings from the 2001 Census in England, Wales and Northern Ireland. Martin then worked with ONS to create Super Output Areas (SOAs) in 2004, merging OAs to enable the aggregation and publication of a range of local data. Samantha Cockings (Lecturer at Southampton since 2000) has since worked on the optimisation of zone design for specific research applications, including reporting population health rates.

From 2008 to 2010, Cockings led research funded by ESRC **[3. G2]** on further developing methods for updating existing systems of small geographical areas when the underlying population characteristics have changed **[3.3]**, **[3.4]**. Martin has investigated the substantive analytical implications of these automated geographies **[3.5]**. Researchers Andrew Harfoot and Duncan Hornby developed the original AZM software into an open-access version, AZTool. Since 2009 Martin and Cockings have further extended the automated zone design methodology to address the particular statistical disclosure control challenges involved in reporting workplace population statistics **[3.6]**, **[3. G3]**.

3. References to the research

3.1 Martin, D. (1998) Optimizing census geography: the separation of collection and output geographies *International Journal of Geographical Information Science* 12, 673-685 [The first peer reviewed academic journal article on this research]

3.2 Martin, D., Nolan A. and Tranmer, M. (2001) The application of zone design methodology to the 2001 UK Census *Environment and Planning A* 33, 1949-1962

3.3 Cockings, S., Harfoot, A. and Hornby, D. (2009) Towards 2011 output geographies: Exploring the need for, and challenges involved in, maintenance of the 2001 output geographies *Population Trends* 138, 38-49 [Peer-reviewed paper in the journal of the Office for National Statistics]

3.4 Cockings, S., Harfoot, A., Martin, D. and Hornby, D. (2011) Maintaining existing zoning systems using automated zone design techniques: methods for creating the 2011 Census output geographies for England and Wales *Environment and Planning A* 43, 2399-2418

3.5 Shuttleworth, I., Lloyd, D. and Martin, D. (2011) Exploring the implications of changing census output geographies for the measurement of residential segregation: the example of Northern Ireland 1991-2001 *Journal of the Royal Statistical Society A – Statistics in Society* 174, 1-16

3.6 Martin, D., Cockings, S. and Harfoot, A. (2013) Development of a geographical framework for census workplace data *Journal of the Royal Statistical Society Series A – Statistics in Society* 176, 585-602

Grants

3. G1) Martin, D. 1999-2000, "2001 output areas: specification, demonstration and maintenance" £43,499, ESRC [first ESRC award]

3. G2) Cockings, S. 2008-2010, "Towards 2011 output geographies: adapting and evaluating automated zone design methods for maintaining the 2001 output geographies" £97,329, ESRC [most recent ESRC award]

3. G3) Cockings, S. and Martin, D. 2009 to date, "2011 output area and workplace zone development" £123,911, Office for National Statistics

4. Details of the impact

The automated zone design procedure developed at Southampton led to the creation of an entirely new nested hierarchy of small areas known as output areas (OAs) covering England, Wales and Northern Ireland that have become the standard for publication of official statistics including 2001 and 2011 census results and neighbourhood statistics service **[5.1]**, **[5.2]**. They overcome many of the shortcomings of previous approaches and enable geographical comparability between successive censuses for the first time in modern history in England and Wales. Using this research, the minimum set of areas requiring redesign has been analytically identified and resolved, resulting in changes to only 2.6% of OAs for 2011. This stability has wide-ranging benefits for data users, including those involved in resource allocations from central to local government, the planning and delivery of services such as housing, education, health and transport, and business planning and marketing, ultimately benefitting all 56.1 million residents of England and Wales. The 2011 census white paper identified as a "key strength" of the Census its "…ability to generate statistics about very small areas and groups of people (as is necessary to ensure that public policies take account of the needs of local communities when formulating policy)" **[5.3, para 6.7]**.

ONS strongly commended the OA geography in its formal 2011 census evaluation **[5.4]** and following public consultation and reviews from 2007-2010, decided to retain the OA geography as the basis for the 2011 census. In addition, the research team were invited to develop an entirely new system of workplace zones for the reporting of census workplace population data. Using our AZTool software, ONS published 2011 output area boundaries in October 2012 and workplace

zones in January 2013.

Our research has directly informed wider policy relating to area design, through Martin's membership of several key advisory boards including the UK Census Design and Methodology Committee, OA Review Panel and the 2011 Census High-Level Quality Assurance Panel. His expertise has also been called upon to inform evaluation of potential future alternatives to the census, being called to present oral evidence to the House of Commons Treasury Sub-Committee 2008 inquiry "Counting the Population" **[5.5]** and Cabinet Office round table meeting in November 2010. These led directly to the launch and direction of the £25m ONS "Beyond 2011" Programme to evaluate potential alternatives to a 2021 census. The impact of our work is emphasised by the Director General of ONS **[5.6]**.

As well as its direct influence on the reporting of UK official statistics, Southampton's research has on-going benefits for data users. In addition to meeting statistical and geographical requirements for disclosure control and regarding the size and homogeneity of the population in each area, the system of OAs is designed with the users' needs in mind. For instance, OAs are uniformly-sized small areas built from, and strongly aligned to, postcode geographies, reflecting the ways in which many users (from local authorities to private companies) categorise their own data. OAs have formed the basis for entirely new data products and user communities, such as the 2001 OA Classification (OAC) **[5.7]** and its user group, currently comprising 288 organisations across central and local government, businesses and academia. A 2011 version of OAC is now in production.

The Super Output Areas (SOAs) resulting from this research form the basis for aggregation and publication of data in the government's Neighbourhood Statistics Service and are integral to many area-based policy initiatives, for example Communities and Local Government's (CLG) Working Neighbourhoods Fund allocations. In 2004, 2007 and 2010 CLG's Indices of Deprivation **[5.8]** were based on the SOAs, permitting the first direct analysis of changes in deprivation over time for nationally consistent geographical areas.

There has been international use of the research and AZTool software; for instance, it has now been implemented by the Australian Bureau of Statistics **[5.9]** to generate their 2011 Area Sampling Frame, a system of geographical areas to be used for the next five years as the basis for Monthly Population Surveys and Special Social Surveys. By making the AZTool software freely available online, benefits have been extended to wider work on official statistics, geographical aggregation problems and applied areas such as public health. Around 25 research groups in ten countries have downloaded the software and reported their usage, including other national statistical organizations such as Statistics New Zealand **[5.10]**.

Overall, this research has been of unique significance for the production and analysis of UK official statistics, providing the framework for local policy decisions and resource allocation. It has substantial international reach, including adoption and evaluation by overseas statistical agencies.

5. Sources to corroborate the impact

5.1 Neighbourhood Statistics Service http://www.neighbourhood.statistics.gov.uk

5.2 Office for National Statistics (2012) Modification of Output Areas http://www.ons.gov.uk/ons/guide-method/census/2011/census-data/2011-census-prospectus/new-developments-for-2011-census-results/2011-census-geography/modifications-of-output-areas/index.html

5.3 Cabinet Office (2008) Helping to shape tomorrow: The 2011 Census of Population and Housing in England and Wales, CM7513: http://www.official-documents.gov.uk/document/cm75/7513/7513.pdf

5.4 Office for National Statistics (2004) *Census 2001 Review and Evaluation. Census Geography: Evaluation Report* http://www.ons.gov.uk/ons/guide-method/census/census-2001/design-and-conduct/review-and-evaluation/evaluation-reports/geography/evaluation-report.pdf [ONS report

published on web: copy in University of Southampton evidence repository]

5.5 House of Commons Treasury Committee (2008) Counting the Population Volume 1 HC-183 1 London: The Stationery Office, http://www.publications.parliament.uk/pa/cm200708/cmselect/cmtreasy/183/183.pdf

5.6 Corroborating letter from Director General (previously Director), 2011 Census, Office for National Statistics.

5.7 Output Area Classification User Group http://areaclassification.org.uk/

5.8 Communities and Local Government (2010) *The English Indices of Deprivation 2010 - Technical Report* https://www.gov.uk/government/publications/english-indices-of-deprivation-2010-technical-report

5.9 Corroborating letter from Assistant Director, Geography, Australian Bureau of Statistics.

5.10 Ralphs, M. and Ang, L. (2009) Optimised Geographies for Data Reporting: Zone design tools for census output geographies Statistics New Zealand Working Paper No 09–01 http://www.stats.govt.nz/surveys_and_methods/methods/research-papers/working-papers/optimised-geographies-for-data-reporting-09-01.aspx