

Institution: University of Southampton
Unit of Assessment: 10 Mathematical Sciences
Title of case study: 10-12 Making It Count: Improving the Census
<p>1. Summary of the impact</p> <p>Statistical techniques developed at the University of Southampton have transformed the accuracy with which Census data can estimate the UK population's size and characteristics, delivering far-reaching socio-economic impact. The methodologies developed by Southampton have increased the accuracy and availability of the 2011 UK Census data, not only for the Office for National Statistics but for central government, local authorities, the NHS and the private sector, who all use the data as a basis for policy decisions. Preserving the privacy of the UK population, Southampton's work allowed, for the first time, the release of highly localised data, which is used by local authorities to target resources efficiently and meet the demands imposed by the Localism and Equality Acts.</p>
<p>2. Underpinning research</p> <p>Government agencies and local authorities rely on accurate census data to make effective planning decisions. The government white paper, <i>'Helping to shape tomorrow: the 2011 Census of Population and Housing in England and Wales'</i> says that: 'The Census underpins the allocation of billions of pounds in funding for public services, and is the foundation of many economic and social statistics. These in turn influence policy across government and investment decisions in the commercial sector.'</p> <p>In the 1991 Census, an estimated 3.8% of the UK population was missed, causing a significant underestimation of local population sizes and provoking considerable frustration among users. This prompted the Office for National Statistics (ONS) to instigate the One Number Census project in 1997. It asked the University of Southampton to develop new methodologies for survey design and estimation to assess the coverage of the population census and adjust the results accordingly.</p> <p>Professor Ian Diamond (1980-2010) led the initial work with Professor Ray Chambers (1995-2006) and Dr James Brown (PhD student 1996-2000; Lecturer 1999 to 2008; Reader/Professor 2010-2013). Brown examined the design of the 2001 Census Coverage Survey (CCS) and the main estimation strategy for calculating key demographic characteristics from the data. The CCS was an independent survey of 320,000 households conducted after the 2001 Census. The results of the Census were matched against those of the CCS, identifying the number and characteristics of those missed in the Census. It led to the development of a dual-estimation system, which was used to amend the census results to reflect those omitted in the original count. For areas not included in the CCS, Southampton's statistical techniques were applied to estimate the undercount [3.1, 3.2].</p> <p>Chambers led research into producing Local Authority estimates of basic demographic data from the main CCS estimates. He developed a system where missing records were imputed based on recorded Census data to produce a complete set of responses for each household and person rather than just basic demographic characteristics. Chambers' idea was tested by Brown, who led the modelling with the LSE's Fiona Steele [3.3, 3.4]. The UK is unique in having such a database, representing the best estimate of what would have been collected had the Census not been subject to under-enumeration. Brown has gained an international reputation for his research in this area and has been invited to review methods being developed for censuses in Canada, Australia and America.</p> <p>Equally important as the accuracy of Census data is the ability to make available detailed census outputs to users, while at the same time protecting the privacy of individual citizens. Southampton's history of research on statistical disclosure control (SDC) dates back to the 1990s with work led by Professor Chris Skinner (at Southampton until 2011) to assess disclosure risk in Samples of</p>

Impact case study (REF3b)

Anonymised Records. This resulted in an EU project led by Skinner, which developed methodologies and software for the assessment of disclosure risk and application of SDC methods to both microdata and tabular data. Between 2005 and 2007 Dr Natalie Shlomo (2007-2012) worked on protecting Census tabular data without compromising its usability. She developed a software package for the ONS for comparing original with protected Census tables in order to assess optimal SDC methods and parameters [3.5, 3.6]. Further development by Shlomo and Skinner [3.7] in response to ONS requirements resulted in modified SDC assessment methods to ensure that statistical outputs provided maximum value to users while protecting the confidentiality of information concerning individuals.

3. References to the research

Publications:

- 3.1 (*) Brown J., Abbott O. and Smith P. (2011). Design of the 2001 and 2011 census coverage surveys for England and Wales. *Journal of the Royal Statistical Society, Series A*, 174(4), 881-906.
- 3.2 Brown J., Abbott O. and Diamond I. (2006). Dependence in the 2001 one-number census project. *Journal of the Royal Statistical Society, Series A*, 169, 883-902.
- 3.3 Brown J., Buckner L., Diamond I., Chambers R. and Teague A (1999). A methodological strategy for a one number census in the UK. *Journal of the Royal Statistical Society, Series A*, 162, 247-267.
- 3.4 (*) Steele F., Brown J., and Chambers R. (2002). A controlled donor imputation system for a one-number census. *Journal of the Royal Statistical Society, Series A*, 165, 495-522.
- 3.5 Shlomo, N. (2007) Statistical disclosure control methods for census frequency tables. *International Statistical Review*, 75(2), pp. 199-217.
- 3.6 Shlomo, N. and Young, C. (2006). Statistical Disclosure Control Methods Through a Risk-Utility Framework. *Privacy in Statistical Databases. Springer Lecture Notes in Computer Science*, 4302, 68-81.
- 3.7 (*) Shlomo, N. and Skinner, C.J. (2010) Assessing the protection provided by misclassification-based disclosure limitation methods for survey microdata. *Annals of Applied Statistics*, 4(3), 1291-1310.

(*) These references best indicate the quality of the underpinning research.

Grants:

- 3.G1 University of Southampton-ONS research contract: Provision of Research Services in Statistical Methodology, 2005-2010 (£850,000) and 2010-2015 (£700,000).
- 3.G2 EU 5th Framework Programme: Computational Aspects of Statistical Confidentiality (CASC), 2000-2003, Chris Skinner, Principal Investigator at the University of Southampton.
- 3.G3 EU 7th Framework Programme: Data Without Borders, 2010-2014, Natalie Shlomo, Principal Investigator at the University of Southampton (£185,000).

4. Details of the impact

According to the *Helping to Shape Tomorrow* White Paper [5.C1]: 'Government, local authorities, the health service, the education and academic community, commercial business, professional organisations and the public at large all need reliable information on the number and characteristics of people and households if they are to conduct many of their activities effectively.' However, in order for all these end users to reap maximum benefit from the Census, 'the information must be authoritative, accurate and comparable for all parts of the country' and '...should be in such a form, and at varying levels of statistical and geographical detail, to meet the

changing requirements of users, subject to the overriding requirement to protect statistical confidentiality.'

Southampton's research into survey design, estimation and statistical disclosure control (SDC) has been integral to major improvements in the accuracy with which the Census is able to estimate the UK population's size and characteristics, and in allowing that information to be released while preserving the rights to privacy of all UK citizens. The immediate beneficiary of Southampton's research was the ONS, who were able to improve the accuracy, reliability and usefulness of the 2011 Census and also save money by limiting the need for extensive additional surveys. Dr Marie Cruddas, Head of Census Methodology at the ONS, said: '*Southampton's research played a key role in the statistical design, estimation and dissemination strategies for the 2011 Census, enabling the production of the high-quality data that the government needs to develop policies, and to plan and run public services.*' [5.1]

The application of Southampton's dual-system estimation allowed basic demographic data of those missed in the Census to be estimated and the adjustments made. For the 2011 Census, the ONS used the CCS in which Southampton's research was the key component of the overall coverage assessment and adjustment strategy. Brown's involvement was cited prominently in the summary of the 2011 CCS [5.2], published on the ONS website. The objectives for the CCS sample design were to select a robust sample that avoided bias, achieved maximum precision and targeted resources in the most important areas. Brown was a central member of the team working on (a) a re-design of the CCS for the 2011 Census, (b) a development of the main estimation strategy and (c) a reassessment of the adjustment system [5.3, see also 3.1]. An independent review team of the coverage assessment and adjustment methodology [5.4] stated: 'The methods give confidence that the resulting final census population estimates will be better than any other method and will be suitable for use in resource allocation and planning.' Further ONS quality assurance work following the 2011 census also used the CCS, and 'added to ONS' confidence in the 2011 Census' [5.5].

Apart from the ONS, the main beneficiaries of Southampton's research are the users of the Census who now have access to data that is both more accurate and available at narrower local geographies than previously possible. The Census drives the allocation of local services such as schools and transport systems, and thus accurate estimates ensure that these services are placed in the right locations. The provision of local health services is also a key factor; the NHS allocates about £120bn a year to trusts largely based on a range of census-derived population estimates, according to the White Paper [5.C1]. Public Health Wales NHS Trust stated: '*It is crucial that we continue to access this gold standard information at small area level.*'

Southampton's work on statistical disclosure control allowed the release of fine-grained local data while preventing the disclosure of sensitive information about individual respondents. This was of particular importance to local authorities, who were relying on the provision of social indicators right down to neighbourhood level – to fulfil their obligations laid out in the Localism Act 2011. The Localism Act marked the beginning of a shift in power from the central government to local authorities that required decisions on issues like planning and housing to be taken locally. Without the local data provided by the 2011 Census, local authorities would have less quantitative evidence to support their decision-making. Of equal significance is the pressure on local authorities to respond to the demands of the Equality Act 2010 for which they need localised data to minimise discrimination among local populations. The importance to local authorities of more localised Census data was fully documented in the ONS 2012 Report: *Beyond 2011 Public Consultation on User Requirements* [5.C2].

The release of microdata to a vast range of end users, including central and local government departments, and the health service, was facilitated by Southampton's SDC research [5.6]. The ONS assessed three SDC methods through a disclosure framework and software package developed by Southampton researchers, Shlomo and Ph.D. student Young. The dependence of the ONS approach to disclosure risk assessment on Southampton's research is described in their

Impact case study (REF3b)

2010 report on Statistical Disclosure [5.7, see page 6]. Based on this work, the ONS decided on the pre-tabular method of record swapping for protecting the 2011 UK Census tables. According to Jane Naylor of the ONS SDC team 'the research undertaken by Natalie Shlomo was central to the final choice of Statistical Disclosure Control (SDC) method used for the 2011 Census. Her research and the software that she developed (InfoLoss) around measuring disclosure risk and utility were used to produce quantitative evidence for the performance of different SDC methods in the Census context' [5.8]. Later development work at Southampton further enhanced statistical disclosure control. This led to its use by the ONS for downstream processing of census data, enabling the release of more useful microdata to a range of end users including central and local government and the health service.

5. Sources to corroborate the impact**Contextual References:**

5.C1 Government White Paper: *Helping to Shape Tomorrow*,

<http://www.official-documents.gov.uk/document/cm75/7513/7513.asp>

5.C2 ONS 2012 Report: Beyond 2011 Public Consultation on User Requirements

<http://www.ons.gov.uk/ons/about-ons/who-ons-are/programmes-and-projects/beyond-2011/reports-and-publications/beyond-2011-user-requirements-consultation-report.pdf>

Sources to corroborate the impact:

5.1 Head of Census Methodology, ONS.

5.2 ONS 2011: 2011 Census Coverage Survey Summary. <http://www.ons.gov.uk/ons/guide-method/census/2011/census-data/2011-census-data/2011-first-release/first-release--quality-assurance-and-methodology-papers/census-coverage-survey-summary.pdf>

5.3 ONS One Number Census Steering Committee.

5.4 Census 2011: Independent Review of Coverage Assessment, Adjustment and Quality Assurance. <http://www.ons.gov.uk/ons/guide-method/census/2011/the-2011-census/the-2011-census-project/independent-assessments/independent-review-of-coverage-assessment--adjustment-and-quality-assurance/independent-review-final-report.pdf>

5.5 ONS 2013 report: Results from using routinely-collected government information for 2011 Census quality assurance. <http://www.ons.gov.uk/ons/guide-method/census/2011/census-data/2011-census-user-guide/quality-and-methods/quality/quality-assurance/cqs-report.pdf>

5.6 Head of SDC Research Team, Methodology Directorate, ONS.

5.7 ONS 2010 Report: Statistical Disclosure Control for the 2011 UK Census. <http://www.ons.gov.uk/ons/guide-method/census/2011/the-2011-census/producing-and-delivering-data/confidentiality/statistical-disclosure-control-for-the-2011-uk-census.pdf>

5.8 Member of SDC Research Team, Methodology Directorate, ONS.