

# **BEYOND 2011: INDEPENDENT REVIEW OF METHODOLOGY**

**CHRIS SKINNER<sup>1</sup>, JOHN HOLLIS<sup>2</sup> and MIKE MURPHY<sup>3</sup>**

<sup>1</sup>Chris Skinner is Professor of Statistics, London School of Economics

<sup>2</sup>John Hollis is a demographic consultant, previously at the Greater London Authority

<sup>3</sup>Mike Murphy is Professor of Demography, London School of Economics

October 2013

## Table of Contents

Abbreviations used in the report .....	3
Executive Summary.....	4
Recommendations .....	9
1. Introduction .....	12
2. Cost Benefit Analysis .....	14
3. The Rejected Options.....	15
3.1. 10-yearly Census with Short Form for Everyone and Long Form for Sample ..	15
3.2. Administrative Data-Aggregate Data Linkage .....	15
3.3. Administrative Data-Partial Linkage .....	15
3.4. Address Register and Survey .....	15
3.5. Rolling Census .....	16
3.6. Short Form Census and 4% Annual Survey .....	16
3.7. Annual Linkage and 10-yearly Survey – 10% and 40% Options .....	17
3.8. Conclusions about Rejected Options .....	17
4. The Online Census Option .....	18
5. The Administrative Data Option .....	19
5.1 Population Estimation .....	20
5.1.1. Quality Targets .....	20
5.1.2. General Approach.....	20
5.1.3. Matching .....	21
5.1.4. Construction of the Statistical Population Dataset (SPD).....	23
5.1.5. Obtaining Data to Assess Coverage Errors in the SPD .....	24
5.1.6. Estimation .....	26
5.1.7. Definitions of population types .....	30
5.1.8. Small local authorities .....	30
5.1.9. Quality assurance .....	31
5.2. Attribute Estimation .....	31
5.2.1. The Options Contrasted .....	31
5.2.2. The Rolling 4% Survey .....	32
5.2.3. Quality Targets: the Rolling 4% Survey .....	32
5.2.4. Content .....	33
5.2.5. Sample Design .....	34
5.2.6. Sample Frame .....	35

5.2.7. Follow-up Design .....	35
5.2.8. Estimation and Adjustment for Non-response .....	35
5.2.9. Compulsory Nature of the Survey .....	35
5.2.10. Dissemination Policy .....	36
5.2.11. Transition from Survey to Administrative-based Attribute Data .....	36
5.2.12. Household Estimates .....	38
5.2.13. Small Area Estimation for Attributes .....	38
<b>5.3. Other Outputs .....</b>	<b>38</b>
<b>6. Transition Issues and Future Validation .....</b>	<b>39</b>
<b>7. Conclusions .....</b>	<b>41</b>
<b>About the Authors .....</b>	<b>43</b>
<b>Acknowledgements .....</b>	<b>43</b>
<b>References .....</b>	<b>43</b>
<b>Appendix 1. Terms of Reference .....</b>	<b>45</b>
<b>Appendix 2. ONS Quality Standards .....</b>	<b>47</b>

## Abbreviations used in the report

CCS: Census Coverage Survey

DWP: Department of Work and Pensions

EU: European Union

LA: Local Authority

LS: Longitudinal Study

LSOA: Lower Super Output Area

MSOA: Middle Super Output Area

NHS: National Health Service

NS-SEC: National Statistics Socio-economic Classification

OA: Output Area

ONS: Office for National Statistics

ONS-M1, ONS-O2 etc. these are references to Beyond 2011 reports and publications – see  
References at end of this Report

PCS: Population Coverage Survey

RSS: Royal Statistical Society

SPD: Statistical Population Dataset

STM: Short Term Migrants

TTW: Travel To Work

## Executive Summary

Up to 2011, the census has been the main source of information in England and Wales for population estimates and the socio-economic characteristics of the population, particularly at small area level. The Office for National Statistics (ONS) has established a Beyond 2011 Programme to consider alternative options. It has now narrowed its focus to two “front running options”:

- **the online census option:** a modernised census – to be carried out every 10 years as at present, and supplemented by the use of administrative data to take account of population change in intervening years.
- **the administrative data option:** an approach that depends on the re-use of the ‘administrative’ data already held within government to estimate the population, supplemented by a 4% rolling annual survey to help estimate characteristics of the population.

The online census option is an updated version of the traditional census-taking model adapted to contemporary technological developments, whereas the administrative data option probably represents the most substantial change in the production of official statistics for over one and a half centuries. ONS is seeking views on which approach best meets user needs in a public consultation launched in September 2013. It plans to provide its conclusions on the Beyond 2011 Programme to Government in 2014.

This report is the outcome of an independent review of the methodological aspects of the Programme, commissioned by ONS at the recommendation of the Royal Statistical Society. The review has sought to assess the methodological research and evidence reported by ONS as the basis of its evaluation of the alternative options; to identify the main risks with the two front running options and to identify areas where further work is required to mitigate these risks; and to enable a sound assessment of methodological considerations relevant to the decision on how to proceed. Most attention has been given to the administrative data option, since it represents a more radical change in methodology than the online census option.

The terms of reference of the review do not include user consultation, nor the assessment of costs, nor detailed operational issues, nor related work in Scotland and Northern Ireland.

The two front running options emerged from a field of ten options identified and evaluated by the Programme. A wide range of useful research has been conducted on these options. We have not identified any methodological issues which would lead us to come to a different conclusion to the Programme, regarding the choice of the two front running options.

The online census option would represent a natural evolution of the traditional census, drawing on technological innovations and developments in best practice for census taking around the world. About 16% of census returns were completed online in England and Wales in the 2011 Census. The return rate of 54% in the 2011 Canadian Census illustrates what might be achievable. The census would be compulsory, with the remaining addresses, where no online return was made, followed up as in 2011.

From a methodological perspective, the 2011 Census in England and Wales seemed to be at least as good as the 2001 Census, for example in terms of response rates. We are not aware that it experienced any major new problems which imply significant methodological risks for the online

census option. ONS has considerable experience with traditional census methodology and has growing experience of the use of the internet for data collection. We believe it has a sound appreciation of the methodological challenges posed by this option.

The administrative data option represents a much more radical departure from established methodology. It is based upon two somewhat separate sources of data. First, a Statistical Population Dataset (SPD) is to be constructed by matching a number of administrative data sources in order to provide annual *population estimates*, i.e. estimated counts of people by sex and age group for local authorities and other areas. Second, a rolling survey of 4% of the population is to be undertaken to collect data on all the other kinds of variables traditionally collected in the census. This would be used to produce what ONS calls *the attribute estimates*. In the longer run, the two separate sources of data could become more integrated by extending the SPD to include attribute variables from the administrative sources and potentially to reduce reliance on the rolling survey.

Administrative data have been used effectively in many countries, including in census contexts. A key requirement is that there is suitable data sharing legislation between the statistical office and the authorities with control over the administrative systems so that these data will meet the statistical needs over time. We consider that appropriate new legislation is a necessary requirement for proceeding with the administrative data option.

The programme has devoted somewhat more research to the methodology of population estimation since the survey methodology required for attribute estimation is much more established. We discuss these two estimation tasks in turn.

The first challenge in constructing the SPD is how to protect the confidentiality of the data subjects when matching records. A significant methodological innovation of the Programme has been to develop a method of matching records whilst preserving their anonymity. It has been shown that this new matching method can achieve almost as high match rates as an established matching method which uses the original records including personal identifying information.

The Programme used such anonymised matching methods to construct some experimental SPDs from data captured from a number of administrative data sources at the time of Census 2011. Population estimates from these SPDs at the local authority level were then compared to estimates from the census. A key empirical finding of the Programme was that there was a reasonable amount of agreement, with the population estimates for one SPD being within +/- 3.8% of the 2011 Census estimates for 90.8% of local authorities.

Nevertheless, another finding from matching the SPDs against Census 2011 data was that there were considerable coverage errors in the SPDs. These errors included under-coverage, where people recorded in the Census were missing in the SPDs, and over-coverage, where people were included in the SPDs by error (e.g. for emigrants or deaths) or where there were duplications in the SPDs or addresses recorded wrongly. The work on the Programme with SPDs is still at an early stage and further improvements of the SPDs, e.g. using additional sources, may be expected to reduce these coverage errors. However, we are clear in our view that population estimates based solely on SPDs would not be able to meet the quality targets that ONS has set.

To address the problem of coverage errors in the SPD, ONS has proposed that a Population Coverage Survey (PCS) be undertaken. This would be analogous to the Census Coverage Survey undertaken at the time of Census 2011 to address the problem of under-coverage in the Census, estimated then to be about 6%. We agree that it is essential that an exercise of this kind is undertaken.

A limitation of the proposed PCS is that it is primarily suited to assessing under-coverage and appears to have only a limited capability for capturing forms of over-coverage. There are other ways to collect data about over-coverage, but these require following up people identified from administrative data to check whether it is valid to include them in the population. The work under the Programme assumes that this would not be publicly acceptable. This seems to us one of the main limitations of the administrative data option. There are reasons to expect that the degree of over-coverage in the SPDs will be significant, more than in census enumeration in 2011, but it is assumed that the natural way of collecting data about over-coverage, as used in some countries with population registers, will not be available.

The estimation methodology used to combine data from the SPD and the PCS is also critical if appropriate account is to be taken of coverage errors in the SPD. The principal approach considered by the Programme is based on dual system estimation and is analogous to the coverage estimation methodology used in Census 2011 with the SPD replacing the Census and the PCS replacing the Census Coverage Survey. This approach is primarily suited to the treatment of under-coverage. The Programme therefore suggests that strategies for minimising over-coverage in the SPD are adopted. There are two risks here. First, even if over-coverage is reduced there will likely be some unknown residual amount which the estimation methodology will not account for. Second, minimising over-coverage risks increasing under-coverage and the robustness of the dual system estimation approach to departures from its assumptions tends to decline as under-coverage increases and this could risk unanticipated errors.

A further risk with the estimation methodology is that it does not adequately take account of errors in matching, especially between the PCS and the SPD. Even relatively small errors could lead to non-negligible biases. It seems possible that the effect of matching errors and over-coverage could be in the same direction, leading to over-estimation of the population, which could be differential between areas.

A significant achievement of the Programme is a simulation study drawing on relevant empirical information, which assesses estimation quality under a range of assumptions. There are some encouraging results from the study, with the relative bias in estimating the population of four large areas generally being less than 1% even when the amount of under- or over-coverage in the SPD was as high as 20%. Nevertheless, the study does not take full account of over-coverage nor of matching error and the study needs further extension if it is to provide sufficient evidence that the quality targets set by ONS are likely to be met.

The programme has suggested an alternative estimation approach, again using the PCS but no longer using dual system estimation. This approach appears to be much less dependent upon assumptions about over-coverage, although it does make other assumptions, in particular that the PCS does not miss individuals within households. Given our concerns about the effect of over-coverage, we consider that this estimation approach should be explored further and its merits relative to the first method assessed.

A further limitation of the simulation study is that it has only obtained results for four large areas. It is hard to know how different results would be for smaller areas and it is necessary to extend the study to consider smaller areas and to study estimates for sex and age groups in further detail.

We now turn to the second estimation task, attribute estimation using the 4% rolling survey. Here the methodology follows conventional sample survey approaches and so the risks are well understood.

We would anticipate the principal methodological risk facing the rolling survey to be the threat of nonresponse bias. As it stands, without census data, there is little evidence in the programme of possible auxiliary variables which could be used to correct for nonresponse bias. We consider that it is essential that the annual survey be mandatory to ensure reasonable response rates and data quality.

The research into attribute estimation is much less developed than for population estimation and we suggest several areas, such as sampling design, where more research would be useful.

If the administrative data option were to be implemented it would represent a step change in the use of administrative data in population statistics in England and Wales. It would perhaps be a shame for such a step change not to be accompanied by the introduction of one or more attribute variables from administrative sources, in particular income, but the Programme has largely considered this as a matter for future development.

Attribute variables from administrative sources could play two roles. Firstly, they may be useful as auxiliary variables for improved design and estimation from the rolling survey. Second, they may be useful as variables of direct interest themselves. We do suggest further research for each of these purposes. The Programme has not yet provided evidence as to whether the introduction of administrative data for attributes could enable the scale of the 4% rolling survey to be reduced, with consequent savings, without overall loss of data quality.

Our focus in this report has been on population estimation and attribute estimation. The adoption of the administrative data option would have significant implications for other census outputs, such as the 1% Longitudinal Study, and we make some brief comments on these implications.

A key issue for either option relates to the transition to the option from Census 2011 and the question of (a) how to judge the extent to which changes in estimates reflect artefacts of changes in methodology rather than genuine changes and, more generally, how to validate estimates, and (b) what to do if there is evidence that some of the estimates are of unacceptable quality.

For the online census option and attribute estimation from the administrative data option, the main concern is with possible mode effects, associated with either measurement or nonresponse, resulting from the steps to achieve more internet returns. We believe such effects can be considered using relatively established survey methodology.

The greater risks, however, appear to be with population estimation under the administrative data option. We believe that it is essential that ONS establish a methodology which enables the validity of population estimates to be assessed. The programme does propose a 10% coverage survey in 2021

for this purpose and we strongly support this idea, although we think the precise nature of the validation method merits further research. Evidence from such validation could help users to interpret the population estimates in relation to previous estimates and could feed into general methodological development. We consider that the validation exercise should be sufficiently substantial that it provides a means of producing alternative estimates if the population estimates produced by the system are deemed to be of unacceptable quality.

A principal question addressed by our review is whether each of the two front running options provides a methodologically sound basis for replacing Census 2011 methodology, given the quality standards specified by ONS.

Our answer to this question is positive for the **online census option**, subject to it including full follow-up for nonresponse and to appropriate work being undertaken to address possible mode effects. We do not believe that there is yet sufficient evidence to support a positive reply to this question for the **administrative data option**. Key elements of such evidence would consist of:

- establishing an estimation methodology which realistically captures the effect of both over-coverage in the SPD and matching error and, using which, it can be demonstrated that neither of these two problems represents a serious threat to the population estimates meeting the quality standards;
- assessment of the likely accuracy of population estimation by age and sex and for areas smaller than local authorities and ;
- demonstrating that quality standards for population estimates at the national level are likely to be met.

Finally, we should like to commend ONS for some excellent work on the Programme and for carrying through their identification and evaluation of options in such a well-organised way. We have made many recommendations for further work on the Programme, but these overlap quite considerably with proposals for next steps within the Programme documents which we have read and we consider that ONS has a strong understanding of the methodological challenges and risks it faces.

## Recommendations

The following 'A' recommendations identify essential work and evidence that we think is required before it would be possible to conclude that the administrative data option provides a methodologically sound basis for replacing the census.

A1. ONS should assess the level of accuracy with which the administrative data option is likely to estimate population by age and sex and for areas smaller than local authorities, specifically MSOAs and LSOAs. It should provide evidence about the nature of the likely estimation errors, especially for smaller areas, and it should provide justification that the level of accuracy is appropriate.

A2. ONS should demonstrate that quality standards for population estimates at the national level are likely to be met.

A3. ONS should establish an estimation methodology which realistically captures the effect of both over-coverage in the SPD and matching error and, using which, it can be demonstrated that neither of these two problems represents a serious threat to the population estimates meeting the quality standards.

A4. ONS should establish a methodology for checking the validity of the population estimates and for producing alternative estimates when these estimates are deemed unacceptable. We expect that this methodology will require substantial additional data collection.

The following 'B' recommendations identify areas of work which we see as priorities for the further development of the Programme.

B1: Attention should be given to measuring the costs and benefits of the two front running options as precisely as possible.

B2: Given the limits of timescale, cost and expertise within ONS, higher priority should be given to addressing a range of methodological issues within the two front runner options rather than to substantial additional work on less favoured options, although the implications of future research for this wider set of options should be monitored.

B3: ONS is encouraged to explore further what systems need to be developed to show how a modernised census based on internet responses could achieve overall and local response rates at least as high as those in 2011. It should also explore how to maximise quality through the use of the address register in the context of multi-mode data collection.

B4: We encourage further development of matching methods. An urgent aim is to quantify feasible error rates (both false positive and false negative), which are needed to determine what quality standards for estimation are achievable. We also recommend further work aimed at the detection and removal of duplications in the SPD.

B5: The requirements for anonymisation need to be reviewed and the possibility of conducting some research in a safe setting under less stringent anonymisation considered.

B6: Further work on the construction of SPDs using additional sources with the aim of reducing coverage errors, as discussed under next steps in ONS-M7 (section 7), should be a priority.

B7: ONS should provide a more detailed and empirically-based assessment of the relative merits of an address-centric SPD vs. a person-centric SPD, including an assessment of the potential benefits of the former for household-level population estimates, and to establish the methodological basis of a strategy for choosing between the two approaches.

B8: A longitudinal SPD would appear to offer many potential benefits and we urge further consideration of this option.

B9: ONS should investigate sources of change in SPDs over time. It should examine the effect of annual changes in SPDs on changes in population estimates and the relation between the changes in the population estimates and other sources of information on population change, such as relating to migration, preferably by running the two systems in parallel for an extended period of time. (See also recommendation B14.)

B10: ONS should investigate alternative sources of information about over-coverage, beyond that provided by the PCS.

B11: The two estimation approaches should be compared, including by simulation.

B12: ONS should establish a methodology for producing confidence intervals which realistically capture, in particular, the effect of over-coverage in the SPD and matching errors.

B13: The properties of estimation methods should be studied not only for the areas already considered in the simulation study, but also for areas of smaller size.

B14: The implication of estimation error on annual changes in population should be studied. Guidance is needed on the size of PCS sample needed to estimate annual change with different levels of precision (extending the work in section 8.3, ONS-M8).

B15: Consideration should be given to what guidance to provide users regarding the interpretation and use of population estimates.

B16: All population types that are currently estimated either from the census or via annual estimation should be listed and it should be assessed which can be estimated under the administrative data option, which could continue to be available by other means and whether any definitions will have to be amended.

B17: Consideration should be given to ensuring comparable quality of estimates across local authorities of different size. Special treatment is likely to be required for the two extreme outliers - City of London and Isles of Scilly UA. The design of the coverage survey should take account of this need.

B18: ONS must test which questions can reasonably be asked on a 4% internet based mandatory survey with expectation of usable answers at all geographic levels (see also Section 5.2.9).

B19: Estimation methods, which make use of the population estimates or other auxiliary information from the SPD or administrative data, need to be investigated for improving quality of attribute data, including both improved precision and for the purpose of bias adjustment for nonresponse.

B20: ONS should consider the best way to disseminate statistics such that there is not a confusing over-supply of different statistics for an LA or MSOA relating to the same 'central year'.

B21: ONS should develop systems that can (a) prepare key OA variables (even with very high CVs) based upon five years data consistent with LSOA data and (b) enable a safe-setting whereby OA variables may be combined to get user-defined geographies.

B22: An investigation should be undertaken into the possibility of constructing one or more auxiliary variables from administrative sources, which would provide useful adjustment for nonresponse bias in the attribute survey.

B23: More case studies should be undertaken, each on a single promising attribute variable which could be obtained from administrative data. The impact of quality issues, such as non-comparability across sources, on alternative analyses using the variable should be examined. For census variables, comparability with previous census estimates should be considered.

B24: In any preparation work for legislation on data sharing, consideration should be given to whether the legislation can be framed in ways which would strengthen the harmonization of attribute variables across administrative sources, with a view to improving their suitability for statistical purposes.

B25: More research should be undertaken on obtaining consistent and comparable household information over time with the administrative data option.

## 1. Introduction

**Background:** The 2011 Census has provided a huge range of statistics about population and housing in England and Wales, both at the national level and for very small geographic areas and population groups. The outputs from the census are used in a multitude of ways inside and outside the public sector. A census has taken place in Great Britain since 1801 and in a broadly similar way every 10 years since 1841, with the exception of 1941. Such census taking also has a long history in many other countries around the world, although a variety of alternative approaches to producing similar statistics also now takes place, especially in countries where there are population registers (Ralphs and Tutton, 2011). In the 2010 census round, apart from the UK only three relatively small countries among the EU-15 members, Ireland, Greece, and Luxembourg, undertook the traditional model (Economic Commission for Europe, 2012). Alternative approaches have received increasing attention in recent years by countries undertaking traditional census data collection for a variety of reasons, including: opportunities afforded by improvements in technology and by alternative data sources; the costs of census taking; rising rates of nonresponse by the public to such forms of data collection; the burden on the public in providing census information; and public concerns about confidentiality.

**Administrative Data:** Administrative data have been useful in the past and have many attractions for use in the future. The marginal cost of obtaining statistics from data already collected primarily for administrative purposes is low. For example, vital registration data on births, deaths and marriage are primary sources of information in these areas. These data cover the whole population (if the event occurs in England and Wales) and are used for a variety of purposes. They may also be combined with other data to provide in-depth analysis of wider socio-economic trends and the Administrative Data Taskforce (2012) considered that “the UK has the opportunity to be a world leader in research using de-identified administrative data routinely collected by government”. There is increasing use of administrative data in census work – or as a replacement for it - around the world. Jeskanen-Sundström (2008) reckoned that twenty countries use or planned to use administrative data in the recent population census round, although these uses do all seem to depend on some form of population register, which is unavailable in England and Wales.

While there are clear opportunities for the greater use of administrative data in official statistics, there are also serious methodological challenges. The variables recorded may not fit the statistical purpose and may be vulnerable to policy changes; people included in the administrative system may differ from the populations of statistical interest; data quality problems may arise e.g. from lags in the systems or from missing data.

**The Beyond 2011 Programme:** Up to 2011, the census has been the main source of information in England and Wales for population estimates and the socio-economic characteristics of the population, particularly at small area level. The Office for National Statistics (ONS) has established a Beyond 2011 Programme to consider alternative options for the production of population and small area socio-demographic statistics for England and Wales. The ONS reported in July 2013 (ONS-P3) that it had narrowed down its focus to two “front running options”:

- “the **online census option**: a modernised census – to be carried out every 10 years as at present, and supplemented by the use of administrative data to take account of population change in intervening years”;

- “the **administrative data option** – an approach that depends on the re-use of the ‘administrative’ data already held within government to estimate the population, supplemented by a 4% rolling annual survey to help estimate for characteristics of the population.”

The online census option is an updated version of the traditional census-taking model adapted to contemporary technological developments, whereas the administrative data option probably represents the most substantial change in the production of official statistics for over one and a half centuries. ONS is seeking views on which approach best meets user needs in a public consultation launched in September 2013. It plans to provide its conclusions on the Beyond 2011 Programme to Government in 2014. Henceforth, we shall refer to the two options above as the ‘online census option’ and the ‘administrative data option’ and to the Beyond 2011 Programme as the **Programme**.

**The Review:** This report arises as part of the quality assurance of the methodological aspects of the Programme. The ONS decided to commission an independent review of this methodological work and asked the Royal Statistical Society (RSS) to recommend an individual to lead the review. The individual was to have appropriate statistical expertise and not to have been involved in the work of the Programme. Chris Skinner was invited by the President of the RSS to take on this role in June 2013. He then asked Mike Murphy and John Hollis to join him to constitute a panel covering the range of experience and expertise specified as desirable in the Terms of Reference. This report reflects the panel’s work undertaking the review between June and September 2013.

**Terms of Reference:** The Terms of Reference of this review are reproduced in Appendix 1. In brief, the panel is asked to review the methodological research and evidence reported by ONS as the basis of its identification and evaluation of the alternative options, to identify the main risks with the ‘emerging view’ and to identify areas where further work or evidence is required to mitigate these risks and to enable a sound assessment of methodological considerations relevant to the decision on whether to proceed with the emerging view. We take the ‘emerging view’ to refer to the two front running options above. We have focussed most of our attention on the risks and methodological challenges arising from the ‘administrative data option’, since it represents a more radical change in methodology than the ‘online census option’. Areas out of scope in this review include: detailed operational issues; ONS work engaging with stakeholders and researching public acceptability; and related work in Scotland and Northern Ireland.

**Review process:** A core task for the panel in the review process was to read the many reports and publications produced by the Programme. The principal ones used are listed in the references at the end of this report together with the Programme website address where they may be obtained. In addition to reviewing these documents, the panel had three meetings with ONS staff, giving ONS the opportunity to expand on some issues and the review team the opportunity to seek clarification. It was not part of the panel’s Terms of Reference to consult users. Nevertheless, the panel did take efforts to ensure it was aware of a range of user needs in order that it did not overlook any significant risks of the emerging view and was able to draw on John Hollis’s work assisting ONS in the preparation of sections of the *Summary of the uses of census information (ONS-C2)*. The panel also occasionally consulted experts outside ONS on specific methodological issues.

**Structure of Report:** The Report is organised as follows: Section 2 offers some brief remarks about cost benefit analysis issues, relevant to final decisions about the options; Section 3 provides a brief

review of the now-downplayed alternative options and Section 4 considers the online census option. Section 5 is the core of the Report and considers the administrative data option. It is divided into two main parts: **population estimation**, that is the estimation of population numbers in local authorities and other areas, possibly by sex and age group, and **attribute estimation**, that is the estimation of quantities derived from attribute variables, such as economic activity or occupation. The separate treatment of these two topics is fundamental to the Programme's research and to the discussion in this report. Sections 6 and 7 discuss some transition and longer term issues and provide some conclusions about the administrative data option.

**Recommendations:** The report contains a number of recommendations to ONS. These are classified as either A or B. The four A recommendations identify essential work and evidence that we think is required before it would be possible to conclude that the administrative data option provides a methodologically sound basis for replacing the census. The larger number of B recommendations identify areas of work which we see as priorities for the further development of the Programme.

## **2. Cost Benefit Analysis**

The information available from these two alternative approaches may have different benefits for users of the two main types of data, population estimates and attribute data. The final choice will be between methods with different costs and benefits, subject to them being identified as robust, acceptable and methodologically sound, and it is the latter issue that is the focus of this review. It is not possible to identify the benefits of the alternative types of data that will be available from these different collection methods. Users of current census data are diverse and census data may form only a part of their final outputs. Many end users may not even be aware that they are, in fact, using census data. New methods of data collection may make new types of data available in timely fashion, but it is difficult to establish what the benefits of such data will be since potential users may not even be known at this stage. One of the purposes of the current consultation exercise taking place between September and December 2013 is to try to firm up some of the estimates of benefits. These data are not available to this review and do not inform the review process.

For some of the Beyond 2011 options, where one option is subsumed in another, it is possible to make estimates of likely cost differentials. With more distinct options, such as the two front runners, only approximate cost estimates can be made at present. It is not feasible to undertake a full cost benefit analysis. We recognise that costs and benefits will play a major role in the final decision and attention should be given to identify and to measure these as precisely as possible. Costs inevitably will be uncertain, e.g. because of unknown rates of nonresponse, and some assessment of the risk of higher costs is needed.

There are a number of secondary issues that might be regarded as methodological apart from the production of population estimates and attribute data where the benefits could be quantified. For example, all currently available census data are potentially available under the online census option, whereas the administrative data option would probably not be able to provide the most detailed information such as on travel to work. On the other hand, the administrative data option could potentially provide small area data about income which would be welcomed by many users.

**Recommendation B1:** Attention should be given to measuring the costs and benefits of the two front running options as precisely as possible.

### **3. The Rejected Options**

The decision process which arrived at the two front running options in section 1 involved a series of other options being rejected, or at least set aside. A short list of eight options was identified during the early stages of the Programme (ONS-M3). Four of these options were eventually rejected, whilst one was expanded to three variants, so that six options were taken forward at that stage (ONS-M3, ONS-P2). These six options were eventually narrowed down to the two front running options (ONS-P3). The four options rejected at this later stage together with the four rejected earlier imply eight options rejected in total and we now comment on each of these.

#### ***3.1 10-yearly Census with Short Form for Everyone and Long Form for a Sample***

This option was rejected since it offered minimal savings but considerably less detailed small-area data relative to a full 10-yearly census (ONS-M3, section 3.3). If cost is a key driver, it is understandable that this option was not considered in detail.

#### ***3.2 Administrative Data – Aggregate Data Linkage***

This option was researched in some detail (ONS-R3). As the potential of administrative data was a key theme of the Programme, it seems appropriate that different ways in which it could be used have been explored. One of the key difficulties in using administrative data for statistical purposes is how to protect confidentiality. Methods which make use of administrative data only through statistical aggregates, as in this option, are very appealing in this respect. The first approach considered of using administrative data aggregates as auxiliary variables in estimation based on coverage surveys seems natural and there are overlaps with the rolling census option research (ONS-R2a). The second approach considered of using modelling without coverage surveys seems more challenging. Whether such methods will provide sufficient estimation accuracy is an empirical question and the negative findings of the research are plausible, although we have not examined this research in detail. Assuming that ONS continues to undertake further research on population estimation using administrative data, it may be worth revisiting this research if and when richer administrative sources become available.

#### ***3.3 Administrative Data – Partial Linkage***

The only account we have of this option is contained in ONS-M2, where it is described as a series of options in which record linkage takes place between one or more administrative files and a coverage survey. It is stated in ONS-M3 that this option and the ‘administrative data option’ are variants of a single approach and we have followed this view by treating this option as subsumed within the ‘administrative data option’.

#### ***3.4 Address Register and Survey***

The title of this option is taken from ONS-M3. It is not described in detail in the papers we have seen nor is the reason for rejecting it clear in ONS-M3 (unlike the other 7 options). As we understand it, drawing on Teague and Swan (2011), this option seems to be similar either to the rolling census option (see 3.5), where an address register acts as a sampling frame, or to the ‘administrative data option’, where the address register may be used to construct the statistical population database

(SPD), which is the basis of population estimates, as well as to serve as a sampling frame for the annual survey. We discuss the rejection of the former option below.

### **3.5 Rolling Census**

A version of this option is currently used in France and the experience gained there might be deemed to lower the methodological risks of this option relative to the 'administrative data option'.

The information collected by a rolling census on attributes might be viewed as similar to that collected from the survey element of the 'administrative data option' (in line with Table 12 of ONS-O2). The discussion of the rolling census option in ONS-O1 assumes between a 4% and a 10% sampling fraction, whereas the 'administrative data option' assumes a 4% fraction. If these fractions are equated, the two options should provide information about attributes to a similar level of accuracy.

The bigger differences between the two options relate to population estimation. However, we have not found the evidence presented by the Programme to provide a clear-cut case favouring one option uniformly over the other. For example, there is some evidence in favour of the administrative data option at the local authority level in ONS-O2. On the other hand it seems plausible that a large component of the estimation error for the rolling census option at the local authority level comes from random sampling error and that this will be much reduced when aggregating to the national level, whereas this will not apply in the same way to the administrative data option.

One reason given in ONS-P3 for rejecting the rolling census option is that it would only offer 'minimal savings' compared to the online census option. One difficulty with the rolling census option is the lack of flexibility from only having one key 'design parameter', the sampling fraction, as a means of controlling costs. If, for example, considerations of the needs for attribute estimates suggested that it was acceptable to reduce the sampling fraction from 10% to 4%, this could lead to an unacceptable loss in quality of the population estimates. The same effect would not occur for the administrative data option.

A final benefit of the rolling census option compared to the online census option is that its greater use of a permanent dedicated staff rather than staff recruited temporarily each 10 years may be expected to lead to higher data quality.

On balance, we agree with ONS that the rolling census does not appear to have clearly-demonstrated superior likely performance to some of the other options while being less flexible in design.

### **3.6 Short Form Census and 4% Annual Survey**

A version of this option is currently used in the USA and the experience gained there might be deemed to lower the methodological risks of this option relative to the 'administrative data option'.

This option is more flexible than the rolling census option in the sense that the sample size of the annual survey can be adjusted without affecting population estimation which is based on the census data.

A key concern reported in ONS-P3 is that this option could lead to a significant increase in cost compared to the 'online census option'. As noted in ONS-P1 (section 5.5.2), this option also places the greatest burden on the public. It is stated in ONS-P3 (section 2.1) that 'there seems little justification to make this change away from our existing approach' and we assume this is why this option has received relatively little research and attention in the papers produced by ONS. We have no further methodological issues to raise regarding ONS's treatment of this option, especially since the two individual components are based on well-established methods.

### **3.7 Annual Linkage and 10-yearly Survey – 10% and 40% Options**

Two options are covered by this heading, corresponding to either a 10% or a 40% 10-yearly survey.

We see no difference in population estimation between these options and the administrative data option, in line with Table 8 in ONS-O2. The key difference is that the administrative data option produces attribute estimates annually rather than 10-yearly.

One advantage of the annual survey, as noted for the rolling census, is that the use of a permanent dedicated staff rather than staff recruited temporarily each 10 years may be expected to lead to higher data quality. However, the primary criterion for choosing between an annual survey and a 10-yearly survey would seem to be what best meets user needs. That is not for us to judge but we have no methodological objection to the argument that the annual survey is likely to prove preferable.

### **3.8 Conclusions about Rejected Options**

- we consider that the Programme has been effective in exploring and identifying possible alternative options to the current census approach;
- we have not identified any other alternative options which we believe should have been considered and could not be subsumed within the options which were considered;
- we consider that some valuable research has been undertaken on the eight rejected options above, especially on the 'administrative data - aggregate data linkage' option (ONS-R3) and the 'rolling census' option (ONS-R2a and ONS-R4);
- we consider that papers ONS-M3, ONS-O1, ONS-P2, ONS-M5, ONS-O2 and ONS-P3 have communicated clearly the Programme's process of decision making, in rejecting eight of the options above for detailed investigation. In particular, these papers are clear about the criteria and quality standards used in this decision making;
- we have not identified any major methodological issues which would lead us to come to a different conclusion to the Programme, that is to end up with other than the two front running options of section 1. Nevertheless, we emphasize that the choice between options depends critically upon user needs and upon costs and it has not been within our remit to assess these needs, nor the specific quality standards employed, nor the cost constraints.

*Recommendation B2:* Given the limits of timescale, cost and expertise within ONS, higher priority should be given to addressing a range of methodological issues within the two front runner options rather than to substantial additional work on less favoured options, although the implications of future research for this wider set of options should be monitored.

#### 4. The Online Census Option

The 'online census option' is one of ONS's two front running options (see section 1). We understand that it would have a format broadly similar to the 2011 Census, and be expected to take place in 2021, following the usual 10 year interval, but with various forms of modernisation, drawing on technological innovations and developments in best practice for census taking around the world.

A key form of modernisation, as conveyed in the option's name, would be modifications to the data collection process with a view to achieving greater completion of census returns via the internet. About 16% of census returns were completed online in England and Wales in the 2011 Census (ONS, 2012a). There are various ways in which a higher rate of return might be promoted, for example Statistics Canada (2012) reports on efforts which led to a 54% internet response in the 2011 Canadian Census.

A basic assumption in our assessment of this option is that all cases where online responses are not received will be followed up fully, using a similar type of approach and level of investment in follow-up data collection to Census 2011. Were the level of follow-up to be scaled back significantly to make savings, we would expect our assessment of the risks of this option to be quite different.

This option is described as being 'supplemented by the use of administrative data to take account of population change in intervening years' (ONS-P3). We understand from ONS that this does not imply an administrative data element to the census option directly, but there may be increased use of administrative data in the production of intercensal mid-year population estimates. We recommend that, were the administrative data option not to be pursued, consideration should still be given to using Statistical Population Datasets (see section 5) for this purpose.

The Programme has understandably devoted relatively little time to this option, since it does not involve any radical change to the approach adopted in the 2011 Census and thus the risks and methodological challenges are well understood.

Although it has not been in our remit to examine the 2011 Census, we are not aware it experienced any major new methodological problems which imply significant methodological risks for the 'online census option', were it to take place in 2021. Indeed, our impression, from a methodological perspective, is that the quality of the 2011 Census was at least as good as the 2001 Census. For example, ONS (2012b) reports that the person response rates in England and Wales were the same, 94%, in 2001 and 2011 and that the minimum local authority response rate increased from 64% in 2001 to 82% in 2011. Moreover, we are not aware of any problem in 2011 comparable to the under-enumeration in Manchester and Westminster in the 2001 Census reported in ONS (2004).

The success of the 2011 Census was based on having developed a good address list, rapid form tracking leading to targeted follow up to areas of high initial non-response and strong community support. A modernised Census based on Internet responses would have to build on that success and meet the challenges of understanding issues raised by a high level of on-line responses. For example within household under-enumeration may be more of an issue if several individual household members each respond on-line for themselves and more individuals may find it difficult to nominate

a particular usual residence if they stay short-term at multiple addresses. Communal establishments would need a separate approach, much as in 2011.

The address list would have to be as good as what was developed for 2011 and would ideally be able to be used after the census. There is a legal hurdle to be overcome but, if this were the case, users would find it much more useful if addresses could be enhanced with those attributes found by the census (e.g. type, tenure, rooms, bedrooms, central heating) that are less likely to change over time.

The content of the questionnaire may require a few alterations to that used in 2011, for example to accommodate same sex marriage and, possibly, the raised state retirement age.

It seems likely that there would need to be follow-up of initial non-response by enumerators at the doorstep. The question of whether these enumerators would merely promote the return of the census form or would key some, or all, variables on a tablet would need to be considered. Those persons without internet access may need special help – possibly including drop-in centres organised within the community.

It is assumed that a Census Coverage Survey (CCS) and Quality Survey would be at the doorstep with tablets. The CCS would require similar targeting and independence as in 2011.

In 2011 there was already a link to administrative data (NHS/Schools Census/DWP) for creation of feasible ranges for census estimates in the quality assurance process. It is anticipated that with the additional experience ONS is gaining on the use of many more administrative sources that the use of such data would be enhanced.

Most of the methodological issues arising from the transition from a largely paper-based to internet-based enquiry arise for the two alternative options, although their emphasis may differ. For both the online census and the administrative data options, a comprehensive and accurate register of dwellings and communal establishments is needed; only at one time point every 10 years for the former but continuously for the latter.

Given the absence of methodological risks of the online census option, it seems entirely reasonable to consider it in a methodological context as a benchmark against which to assess alternative options which depart more radically from the 2011 Census approach.

*Recommendation B3:* ONS is encouraged to explore further what systems need to be developed to show how a modernised census based on internet responses could achieve overall and local response rates at least as high as those in 2011. It should also explore how to maximise quality through the use of the address register in the context of multi-mode data collection.

## **5. The Administrative Data Option**

The Programme has used administrative data captured at the time of Census 2011 as a basis for research. Its discussion of this option generally takes administrative data as given. More attention could be given to the question of how administrative data systems could be adapted to improve the statistical performance of this option. It is essential that this question is addressed when formulating data sharing legislation.

The Programme has made a strong distinction between population estimation, which draws especially on administrative data organised into a Statistical Population Dataset in this option, and attribute estimation, drawing especially on a rolling survey in which the attribute variables are measured. We follow this distinction by considering population estimation in section 5.1, followed by attribute estimation in section 5.2, with a short final section on other outputs.

## **5.1 Population Estimation**

### **5.1.1 Quality Targets**

Targets (e.g. all LAs have 95% confidence interval of +/- 3.8% or better) have been set for national and LA level population estimates based on performance of the current system (see Appendix 2). It is essential also to ensure that population estimates are produced, at appropriate levels of accuracy, for age/gender groups down to single years of age as well as for middle super output areas (MSOAs) and lower super output areas (LSOAs). While overall population totals are important it is vital that all users of local authority (LA) level data can view age groups on which service delivery is based, with as much confidence as at present. It is critical that the quality of each LA estimate is similar to the quality of all others as competitive resource allocation may be made on the basis of the estimates or projections developed from them. At present users of mid-year estimates tend to use the numbers 'as given' but in future it will be necessary to understand the quality of those estimates in order to assess the risk of making decisions based on inadequate data.

The targets at the national and LA level are expressed in terms of summary performance. There is also a need for a lot more evidence about the nature of the estimation errors. For example, we need to know if there are certain types of areas, inner urban for example, where estimation errors are likely to be systematically worse, i.e. they are 'harder to count', and we need evidence about the likely directions of the biases. Such evidence is needed especially for the more detailed estimates for smaller areas and for age/gender groups.

*Recommendation A1.* ONS should assess the level of accuracy with which the administrative data option is likely to estimate population by age and sex and for areas smaller than local authorities, specifically MSOAs and LSOAs. It should provide evidence about the nature of the likely estimation errors, especially for smaller areas, and it should provide justification that the level of accuracy is appropriate.

### **5.1.2 General Approach**

The proposed approach is outlined in Figure 1. Central to the approach is the Statistical Population Dataset (SPD) which is constructed from two or more administrative datasets using a set of 'rules' to determine, based on matching, which records to include from each dataset. The objective is to include a single record in the SPD for each individual in the population and no records for individuals outside the population. In practice, it is recognized that this objective will not be achieved perfectly and a Population Coverage Survey (PCS) will be undertaken to assess errors of coverage. The estimation method adjusts the population estimates coming from the SPD for coverage error using the PCS. The approach is similar to that adopted in the 2011 Census, with the SPD replacing the census and the PCS replacing the Census Coverage Survey (CCS).



**Matching methods:** Some valuable work on matching methods has been undertaken under the Programme. There has been some useful research on how to use automatic matching methods to avoid input from clerical staff. There has been some particularly innovative research designed to protect the confidentiality of the sources, one of the harder challenges raised by the administrative data context. Smart approaches have been developed to match anonymised records, while retaining good match rates by clever pre-processing methods prior to anonymisation.

**Anonymisation:** The necessity of the form of anonymisation process proposed in ONS-M9 is not discussed in any detail and an explicit discussion of the options for anonymisation seems needed. For example, even if the proposed anonymisation process is adopted in more routine work, it may be valuable if matching of the original data can be undertaken in an appropriately safe setting for some specific quality assessment research. For example, there may be some kinds of duplication or address misclassification which can be identified much more effectively when matching the original data and such research may be helpful in e.g. the determination of rules for combining data sources. It would be unfortunate if this possibility has to be ruled out.

**Extensions to methods:** This initial research is promising and we urge ONS to continue it. The field of matching is an active one around the world today (e.g. Christen, 2012) and it is important to consider whether error rates can be improved, especially since errors in matching the PCS to the SPD may have serious biasing consequences for estimation (see section 5.1.6). Various sensible lines of further development are set out in ONS-M9 (sections 5, 6). It certainly seems desirable to explore the use of further matching information, such as from addresses and from historical data. We understand the preference for methods which do not require clerical review but it seems rather early to rule this option out.

**Duplications:** We recommend more research into the removal of duplications and, more generally, the possibility of correcting erroneous information about location. There is evidence from previous census research, especially in the US, that duplications can be a serious problem and are not easy to identify (Hogan, 2003; Mulry, 2007). Relevant software has been developed, e.g. BigMatch (US Census Bureau) and Monster Match (Statistics Canada). Duplications may arise quite differently in an SPD, whether within an administrative source or in more than one administrative source where a link has failed to be established. However, it still seems desirable to research the feasibility of using matching methods to identify duplicates as well as the possibility of correcting address classification in cases where e.g. two records match in all respects except address.

**Differential error rates:** Some research exploring differences in matching rates between different groups, e.g. by age group or ethnicity, may prove useful not only in method development but also in the study of possible differential biases at the estimation stage (as in section 5.1.6).

**Design of the PCS to facilitate matching:** We agree with ONS-M9 that it would be sensible to consider different ways of collecting information in the PCS to improve linkage to administrative data bases. Experimenting with alternative modes of survey data collection may be worthwhile.

**Recommendation B4:** We encourage further development of matching methods. An urgent aim is to quantify feasible error rates (both false positive and false negative), which are needed to determine what quality standards for estimation are achievable. We also recommend further work aimed at the detection and removal of duplications in the SPD.

*Recommendation B5:* The requirements for anonymisation need to be reviewed and the possibility of conducting some research in a safe setting under less stringent anonymisation considered.

#### **5.1.4 Construction of the Statistical Population Dataset (SPD)**

**Source reports:** ONS has conducted some valuable research evaluating a number of alternative administrative data sources which might feed into the construction of the SPD (ONS-S1, S2, S3, S4 and S5). These reviews provide useful details of the administrative contexts from which the sources are derived and much information about how coverage errors could arise.

**SPD-Census comparisons:** Some of the key empirical findings of the Programme are reported in ONS-M7, where a number of SPDs are constructed using administrative data captured at the time of Census 2011 and estimates of population for local authorities derived directly from the SPDs are compared with estimates from Census 2011. The results suggest that the approach using SPDs holds some promise, with the population estimates for one SPD (SPD5) being within +/- 3.8% of the 2011 Census estimates for 90.8% of local authorities. This is before any coverage adjustment is applied.

**Rules and SPD development:** We support further research developing SPDs along the lines outlined in section 7 of ONS-M7. For example, it makes sense to explore the potential to use death registrations to remove people from the database. The methodology underlying the rules for combining administrative sources and for resolving differences in addresses or ages, for example, would benefit from a more systematic basis, as in Harper and Mayhew (2012), for example. It will also need to be made clear to all users. To use a 50:50 split for two separate addresses in two different sources may be best in many cases – but in others there may be more recent supporting activity data. For example the NHS will capture a re-registration date or a first registration date, such as Flag 4s for new registrants from overseas, and also has episode statistics of in- and out-patient treatment that contain dates of activity. DWP data will have dates of a change of employer and of starting to claim particular benefits. These additional sources may be able to determine a most likely residence address at the relevant date of the population estimate. The strategy of using rules to minimise over-coverage in the construction of the SPD will be considered in sections 5.1.5 and 5.1.6.

**Address-centric SPD:** Research on the construction of the SPD has so far focussed on matching at the person level. We do think that more research is needed into the possibility of constructing an address-centric SPD. This could have a number of possible benefits, including (i) reducing coverage errors in the SPD through improved matching ability, (ii) potential to use with weighted estimation approach (see section 5.1.6), (iii) provision of household-level population estimates and (iv) integration with address register with consequent benefits for the sampling frame of the continuous survey. There may be lessons from experience in other countries, in particular from Germany, where the new register-based census seems to draw on extensive matching work at the address level between register-based administrative sources as well as a census of building and housing. Especially given the importance of small-area household estimates, the methodological pros and cons of a person-centric and an address-centric SPD need to be reviewed and a strategy established for choosing between them.

**Longitudinal SPD:** A longitudinal SPD would appear to offer many potential benefits, including improved matching quality and the potential for richer population estimates, such as for migration

(ONS-M8, Section 3.3) and improvement of sample design of the 4% annual survey. We urge further consideration of this option.

**Change over time:** Users are very keen to get reliable, consistent estimates of population change, which can be broken down into natural change and migration, for use in analysis and modelling. Natural change will still be available from vital statistics and will, presumably, be used in quality assurance. It is assumed that migration will continue to be estimated by ONS from a number of sources, mainly the NHS Central Register and the International Passenger Survey, with the potential of data from e-borders. One significant risk of the administrative data option is the potential effect on population estimates of changes in the SPDs over time, resulting from changes in the administrative systems rather than from genuine changes in the population. It seems feasible that the quality of annual estimates of change, even at LA level, could be poor even if the quality of the estimates is good. Annual change in the estimates may not match the combination of migration and vital registration. It seems essential to assess the stability of SPDs over time, as proposed in ONS-M7 (section 7). Changes in SPD-based estimates need to be compared with changes in mid-year estimates (based on established ONS methodology), and the differences between the outcomes of the two approaches understood. This work seems vital if the population estimates from the administrative data option are to replace the mid-year estimates.

*Recommendation B6:* Further work on the construction of SPDs using additional sources with the aim of reducing coverage errors, as discussed under next steps in ONS-M7 (section 7), should be a priority.

*Recommendation B7:* ONS should provide a more detailed and empirically-based assessment of the relative merits of an address-centric SPD vs. a person-centric SPD, including an assessment of the potential benefits of the former for household-level population estimates, and to establish the methodological basis of a strategy for choosing between the two approaches;

*Recommendation B8:* A longitudinal SPD would appear to offer many potential benefits and we urge further consideration of this option.

*Recommendation B9:* ONS should investigate sources of change in SPDs over time. It should examine the effect of annual changes in SPDs on changes in population estimates and the relation between the changes in the population estimates and other sources of information on population change, such as relating to migration, preferably by running the two systems in parallel for an extended period of time. (See also recommendation B14.)

### **5.1.5 Obtaining Data to Assess Coverage Errors in the SPD**

**Need for data to assess coverage error:** There is evidence of considerable coverage error in the SPDs (ONS-M7, section 6). These errors could be assessed in 2011 by comparing the SPD to Census 2011. If the administrative data option is adopted, however, it will not be possible in the future to use the census to undertake such coverage assessments and it will be essential to have some alternative means of undertaking such assessment. This will be required, at the very least, for validating population estimates derived from the SPD (discussed further in section 5.1.9 on quality assurance), but also, given the scale of the under- and over-coverage in potential SPDs revealed by

ONS, as a basis for adjustment of estimates obtained directly from the SPD, as will be discussed in section 5.1.6.

**Approaches to assessing coverage error:** The natural approach to assessing under-coverage in the SPD directly is first to identify a sample of individuals who should be in the reference population and then to check whether they are in the SPD (by record linkage). This is the basis of the approach proposed in ONS-M7 using a Population Coverage Survey (PCS), modelled on the Census Coverage Survey (CCS) used with the 2011 Census. Conversely, the natural approach to assessing over-coverage in the SPD (assuming that duplications have already been addressed – see section 5.1.3) is first to take a sample of records in the SPD and then to follow these up to the recorded addresses to check whether their inclusion in the SPD is valid. Examples of this type of approach to over-coverage have been used in the 1990 and 2000 US Census (Hogan, 1993,2003), the 2000 census in Switzerland (Renaud, 2007) and the population register-based census in Israel in 2008 (Nirel and Glickman, 2009; Ben Hur and Burck, 2013).

**Dependent surveys:** A basic problem for the administrative data option is that such an approach to collecting data to assess over-coverage is deemed infeasible. Thus, it is stated in ONS-M7 (section 3.3) that “it is assumed that a dependent survey (where the sample is drawn directly from either a single administrative data source or a combined set) will not be possible”. This assumption is based upon the judgement that this would not be publicly acceptable, even if enabled by new legislation, for sources such as the NHS Patient Register, a point we return to in the final paragraph of this section.

The same objections to a dependent survey do not apply for a traditional census, as in the USA or Switzerland, or for a register-based census as in Israel. Unfortunately, this seems to put ONS in the position of having to measure over-coverage using less information than is available to any other national statistical agency at the same time as having potentially a greater problem of over-coverage than most other agencies have to face. For example, erroneous inclusions from SPD records not being removed for emigrants should be a much smaller problem in a traditional census and it seems feasible that lags in changes of address in administrative data bases could lead to more misclassification of address than in a traditional census. Although there can be serious concerns about address misclassification in countries, like Israel, that use population registers in the place of the SPD, there seem no grounds to suppose that over-coverage of this kind will be less of a problem for the proposed SPD than for such population registers, indeed the reverse seems more likely. A crude estimate is that 7% of records in the SPDs considered by ONS so far represent erroneous inclusions (a figure obtained from section 6 of ONS-M7 by subtracting the estimated 6% nonresponse in the Census 2011 from the 13% of administrative records which cannot be linked to the census). The absence of any direct means to assess such potentially large amounts of over-coverage in SPDs in the future seems a major limitation of the administrative data option, although this does depend on the form of estimation method (see section 5.1.6).

**Over-coverage minimization:** One strategy suggested in ONS-M7 and ONS-M8 to address the over-coverage issue and, specifically, the issue of erroneous inclusions is to construct the SPD so that over-coverage is minimized. Thus, the first rule suggested for the construction of the SPD is to include individuals only if they appear in two data sources with the same address with the aim that this ‘should minimise the number of erroneous inclusions’ (ONS-M7, section 3). There must be many

ways to reduce over-coverage in administrative sources, e.g. by removing records where no activity, such as payment of taxes or attendance at benefit offices, is recorded. Indeed, ONS tells us that the Programme aims to use more activity data, such from the Lifetime Labour Market Database, to do this and the Programme is researching ways of using the administrative data to build models to identify and remove such people from the SPDs. Whilst it is clearly desirable to remove records for which the evidence of erroneous inclusion is overwhelming, e.g. when a link has been made with a death registration, the question of whether to remove the record becomes less clear when the evidence is less certain. In this case, there is the potential for generating increased under-coverage, which may lead to other problems at the estimation stage, an issue we return to in section 5.1.6. Moreover, even if many erroneous inclusions can be removed, it is unlikely that all will be and so there will remain the need to check how far the strategy is successful.

***Use of the Population Coverage Survey (PCS):*** The PCS may be usable to estimate some components of over-coverage as well as all under-coverage (ONS-M8). The Programme has focussed on the method of Large et al. (2011), which is designed to assess a component of over-coverage arising from misclassified addresses. The method treats cases in the PCS which are linked to records in the SPD in different areas as evidence of such over-coverage and it is evaluated in the simulation study (see section 5.1.6). A more novel and less researched method is based on the observation (ONS-M8, Sect 6.3) that, if the PCS seeks to fully enumerate an area, those records in the SPD in this area which cannot be matched to the individuals found in the PCS represent either over-coverage in the SPD or nonresponse in the PCS (leaving aside linkage error). It is suggested that if multiple administrative sources are used it may be possible to disentangle the over-coverage and nonresponse processes and hence to estimate the amount of over-coverage in the SPD. A lot more research is needed to establish whether this can be done reliably. We comment further on the role of over-coverage in estimation in section 5.1.6.

***Other sources of data:*** It is not clear to us that dependent surveys should be ruled out entirely. For a start, an SPD could be matched to an address register and addresses not found on the register followed up by a field visit to check at least partially on their validity even if people are not interviewed. Moreover, for some components of an SPD, perhaps electoral register or HESA records on students, it is not clear to us that interviewing a subset of people obtained from such sources would necessarily be either unethical or unacceptable. For example, the electoral register, a public document, has often been used as a sampling frame, although there are some restrictions on its use for commercial purposes. Another possibility, which should not breach individual privacy, would be to visit communal establishments listed as addresses in the SPD to check, first, whether they exist and, second, whether counts of residents in the SPD match counts recorded by the establishments. ONS will be much better placed than us to come up with other ideas of this kind. Such possibilities seem worth exploration.

***Recommendation B10:*** ONS should investigate alternative sources of information about over-coverage, beyond that provided by the PCS.

### **5.1.6 Estimation**

***First estimation method and simulation study:*** The estimation approach which has received most attention follows that for Census 2011 and involves a combination of ratio estimation, dual system estimation, and the over-coverage correction of Large et al. (2011). The adjustment of the

enumerated population counts using the Census Coverage Survey (CCS) in Census 2011 is replaced by an analogous approach where the SPD counts are adjusted using the PCS (ONS-M8). This simulation study provides valuable information. The study considers a wide range of assumptions about nonresponse in the PCS and coverage errors in the SPD. The realism of the study is strengthened by using assumptions about the SPD that have been derived from an exercise in which the NHS Patient Register was matched to 2011 Census and CCS data and other administrative data (we understand that this will be further strengthened using probabilities derived from SPDs). Under the range of assumptions considered in the simulation study, the proposed estimator seems to perform quite well, with the relative bias in estimating the population of four large areas generally being less than 1% even when the amount of under- or over-coverage in the SPD was as high as 20% (ONS M-8, Tables 6 and E2). It was found that the relative bias in five year age groups by sex could be somewhat larger, up to 3% (ONS M-8, Figures 3 and E2) and it is reported that 'at age-sex detail the patterns show that the bias can be quite significant for groups with high over-coverage on the SPD'.

The SPD-based estimates without any coverage adjustment do not meet the national level quality targets in ONS-M7, section 5.1. As acknowledged in ONS-M7, it is necessary to assess whether coverage adjustment will enable the targets to be met, whether by extending the simulation study in ONS-M8 or by some other way.

**Effect of erroneous inclusions:** Although these results might be interpreted as encouraging, the simulation study does make some important assumptions. As noted in ONS-M8 (section 9), two seem of key importance. First, the study only allows for certain kinds of over-coverage, specifically the kind of misclassification of address within broader (local authority) areas which the over-coverage correction of Large et al. (2011) is designed to address. The study does not allow for 'erroneous inclusions' in the SPD, i.e. records which are not part of the population. ONS-M8 acknowledges that erroneous inclusions will not be corrected for by the proposed approach so that the inclusion in the SPD of 1% too many people who should not be there will simply lead to 1% too many people in the population estimate.

**Effect of matching errors:** Secondly, no allowance is made for errors of matching between the PCS and the SPD. We agree that "it is crucial to the Programme to understand how matching errors will impact on the accuracy of population estimates" (ONS-M9, section 7). It would seem that false positive matches will tend to lead to under-estimation of the population (downward bias) and false negative matches will tend to lead to over-estimation (upward bias). The findings in ONS-M9 (section 4.2.2) suggest that the predominant error will be false negatives (values between 0.9% and 3.3% are reported for six local authorities). This suggests that the bias resulting from matching errors will be in the same direction as that resulting from erroneous inclusions in the SPD, i.e. both will tend to lead to overestimates of population. The fact that these errors may have variable effects across areas may be of particular concern. For example, it may be that the degree to which population is over-estimated is greater in urban areas with greater population change.

**Second estimation method:** A quite different approach to estimation is also suggested (ONS-M8, section 6.1), avoiding the need for dual system estimation and the over-coverage correction and their associated assumptions. This alternative method again uses the PCS, but this is no longer required to be independent of the SPD. Indeed, it may be possible to use the SPD to inform the

selection of the PCS sample. An additional assumption, however, is that the PCS provides a correct count of people at each responding address. Assuming again some clustering of the sample, a similar kind of ratio estimation to that used in the first estimation method, treating SPD counts as the auxiliary variable, can be used to estimate all clusters in the population from the sampled clusters. The main novel idea is to estimate the population count within each cluster also using ratio estimation, treating the SPD count of people at an address as the auxiliary variable and the PCS count of people at a responding address as the true count of interest. One could alternatively interpret this approach as using a particular kind of within-cluster imputation for people at addresses which feature in the SPD in the cluster but are missing from the PCS. A potential claimed advantage of this estimation approach is that it is robust to over-coverage in the SPD. The SPD counts are only used as auxiliary variables and so they can be subject to error, so long as they remain well correlated with the address-level counts from the PCS. It is not clear to us, however, that there will be no biasing effects of any kind of over-coverage. This may be true for erroneous inclusions within valid addresses which may be sampled in the PCS. However, it would seem that erroneous inclusions in addresses that are non-existent (or otherwise have zero chance of inclusion in the PCS) may lead to upward bias. Perhaps the main concern with this approach is the assumption that people are correctly counted at responding addresses in the PCS. Another possible concern is that the estimates may have increased variance and this may be an issue for comparisons over time. This approach would seem to attach a greater relative importance to getting a good quality PCS survey vs. a good quality SPD.

This approach is address-based and is most natural if the SPD were address-centric and it may make sense for the address register to feature both in the construction of the SPD and the sampling frame for the PCS. The person-level matching between the SPD and the PCS required for the first estimation method is replaced by address-level matching and this may be an easier task. Given the difficulty that the first estimation method has with handling over-coverage, we do recommend that this alternative estimation method be explored as well, in particular by implementing it in the simulation study and studying its properties relative to the first estimation approach under a range of realistic assumptions. Just as it will be essential to consider the impact of matching errors on the first estimation method, the impact of address-level matching errors on the second method should also be considered.

***Effect of minimising over-coverage in SPD on first estimation method:*** It is necessary to evaluate the relative merits of the estimation methods in conjunction with the rules for constructing the SPD. Thus, we referred in section 5.1.5 to the minimisation of the over-coverage in the SPD if the first estimator is to be used. A different SPD construction strategy might be preferred for the second estimator. It seems possible that, although an over-coverage minimisation strategy might protect against bias from erroneous inclusions in the first estimation method, it may risk an increase of bias from under-coverage as well as an increase in variance. The potential variance inflation effect is acknowledged in ONS-M8 (sections 8.2.3 and 9) and it is noted that “further simulations with more realistic over-coverage patterns will provide better information about the likely variance of estimates”. Earlier work in Brown et al. (1999) suggested that the robustness of the dual system estimation method to assumptions about ‘dependence’ declines as the rate of under-coverage increases, and stated (p.266) that ‘it is vital to obtain a very high CCS response’ to achieve such robustness (here CCS refers to the Census Coverage Survey, equivalent to the PCS in the administrative data option). There is a tension between this effect and the strategy to minimize

over-coverage. Again, further work is needed. The current simulation study seems to make little allowance for such 'dependence' and the robustness of the proposed approach under realistic amounts of under-coverage implied by a possible over-coverage minimisation strategy certainly seems to merit further research.

***Compulsory nature of PCS:*** Following on from the above quotation from Brown et al. (1999), it does seem essential that the PCS be mandatory, alongside the 4% rolling survey (see section 5.2.9) to ensure that adjustment for under-coverage can be undertaken accurately and robustly.

***Estimation for smaller areas:*** At present, the simulation study only provides results for four large areas. It is hard to know how the estimation method will work for areas of smaller size. The first estimation method is based closely upon the estimation method used for the 2011 Census. Whereas there are likely to be some common features of coverage error in the census and the SPD which can be handled by a common method, a major difference would seem likely to be the much greater degree of address misclassification in the SPD. A consequence may be that the first estimation method becomes much more problematic the smaller the area for which the population is being estimated and that the degree of degradation by area size is worse for the SPD than for the census. This implies that it is difficult to predict performance on the basis of ONS's experience of population estimation at different levels from the 2001 and 2011 censuses. This is something that could be looked at in an extended simulation study. However, it would seem to require a lot more work since the current simulation engine seems to make very little allowance for geography. It would be strongly desirable that a further simulation study to address this issue would misclassify addresses in the SPD in a realistic manner based upon empirical evidence.

***Recommendation A2:*** ONS should demonstrate that quality standards for population estimates at the national level are likely to be met.

***Recommendation B11:*** The two estimation approaches should be compared, including by simulation.

***Recommendation A3:*** ONS should establish an estimation methodology which realistically captures the effect of both over-coverage in the SPD and matching error and, using which, it can be demonstrated that neither of these two problems represents a serious threat to the population estimates meeting the quality standards.

***Recommendation B12:*** ONS should establish a methodology for producing confidence intervals which realistically captures, in particular, the effect of over-coverage in the SPD and matching errors.

***Recommendation B13:*** The properties of estimation methods should be studied not only for the areas already considered in the simulation study, but also for areas of smaller size.

***Recommendation B14:*** The implication of estimation error on annual changes in population should be studied. Guidance is needed on the size of PCS sample needed to estimate annual change with different levels of precision (extending the work in section 8.3, ONS-M8).

***Recommendation B15:*** Consideration should be given to what guidance to provide users regarding the interpretation and use of population estimates

### **5.1.7. Definitions of Population Types**

At present, ONS produces estimates for a number of separate population types, either directly from the Census or through the decade by estimation. Should the administrative data option be used for population estimation, close consideration needs to be given to what definitions are appropriate to the methodology:

- a) *Usual residents*: It is necessary to be able to distinguish any communal populations. Difficult communal populations that are included in the current mid-year estimates are armed forces and dependents, prisoners (not all of whom are included as resident at their place of incarceration), boarding pupils and students at their term-time address.
- b) *Short-term migrants (STM, <12 months)*: The 2011 Census provided one set of estimates of persons *intending* to be in the UK for between 3 and 12 months and administrative sources (International Passenger Survey and DWP) are used to estimate local STM populations each year. Some STMs may appear on the administrative sources used in the linking. It will be necessary to find ways of filtering out such people, possibly using NHS Flag 4 indicators and some additional surveying. There are also similar issues with the rolling survey of attributes where STM may be interviewed and will need to be separated out.
- c) *Daytime population*: The census provides workplaces as could an attribute survey. School address of schoolchildren (apart from those at private schools) and student places of education are available in administrative data.
- d) *Armed forces and dependents*. Armed forces in private households may be on duty overseas and should be included in the usual resident count, even though some may not be on administrative sources. Armed forces in communal establishments are not necessarily on local NHS lists. Foreign armed forces and dependents are less likely to be in administrative sources and are in any case very difficult to estimate by usual means.
- e) *Population Present/Visitors*: these categories are probably not possible with administrative sources.

*Recommendation B16*: All population types that are currently estimated either from the census or via annual estimation should be listed and it should be assessed which can be estimated under the administrative data option, which could continue to be available by other means and whether any definitions will have to be amended.

### **5.1.8 Small Local Authorities**

Local Authorities (LAs) are now responsible for their own Local Plans under 'localism'. Population estimates – and the accompanying change analyses – are used as the basis for local population projections by ONS and by individual authorities. With a user requirement for consistent annual data across all LAs – as for the current mid-year estimates – ONS must consider ways to improve the potential quality of data for small LAs.

ONS has used examples based on the (mean) average population size of LAs. At mid-2012 this was 162 thousand. However, the distribution of LA sizes is not regular and 67% of LAs have below average populations. The median LA is Eastleigh with 127 thousand residents. Of the 182 cells of single years of age/gender in Eastleigh, 119 contain fewer than 800 people. (Lacking a quality target for the estimates by age group, 800 is used in this example as it is the size of a sub-population that

may be reliably estimated via the attribute survey in one year.) There are 201 LAs (non-Metropolitan District and non-Unitary) of average size 107 thousand. Of these 95% are below the overall average size (162 thousand) with populations down to 35 thousand, at this level an estimate of 800 is 2.3% of the population rather than 0.5% of the average LA. The estimation methodology needs to be targeted to achieve high quality estimates in all areas. While the rationale for targeting the coverage survey at 'hard-to-count' areas is understood, it should also be targeted with the aim of equalising error levels across LAs of different sizes.

*Recommendation B17:* Consideration should be given to ensuring comparable quality of estimates across local authorities of different size. Special treatment is likely to be required for the two extreme outliers - City of London and Isles of Scilly UA. The design of the coverage survey should take account of this need.

### **5.1.9 Quality Assurance**

There is a need to develop methods for quality assuring the population estimates. For the 2011 Census, administrative data were used to estimate feasible bounds for the census estimates by age and gender in each LA. Administrative data feed more directly into the population estimates in the administrative data option, but it may be possible to use additional administrative data for some checking of estimates. Demographic information could be used in a variety of ways. Birth and death occurrences in the previous 12 months could provide checks. Births by gender can be used as a check on under-1s. Internal and international migration estimates collected in improved ways should still play a role as a comparator for population change even where they may not be used directly in the improvement of a single set of estimates. Currently the mid-year estimates are known to drift between census years and research is on-going to better understand this phenomenon. There will need to be regular checks of 'drift' using as many other sources as possible. In the past the Census has been used as an overall decennial check on migration estimates through the decade. In future the population estimates may have to be checked by reference to migration estimates. Other more regular demographic 'sense' checks (e.g. sex ratios, general and total fertility rates, life expectancies) must be carried out. They may not impact the production of estimates in a single year but will either give added confidence on the estimates or point to potential inaccuracies in some LAs. This topic is returned to in section 6.

## **5.2. Attribute Estimation**

### **5.2.1 The Options Contrasted**

Methodological issues associated with the alternative options for producing population estimates were discussed in section 5.1. In the case of attribute data, the distinction between the two main approaches is more substantial. In the online census option, attribute data have a small marginal cost given that all households and communal establishments will be contacted for basic population estimation purposes. In the case of the administrative data option, the processes for obtaining population estimates and attribute data are largely independent.

Nevertheless, there are fewer methodological issues with the attribute data in the administrative data option since this is substantially based on conventional sample survey approaches, but there are a number of issues specific to a survey closely associated with the rest of the administrative data package. These issues include: the quality targets; possible use of SPD and other administrative data to improve precision of estimates and adjust for non-response; and development of small area estimation methods to obtain additional variables at small area level. To date, much less research has been undertaken on these issues and comments are therefore inevitably less detailed. The longer-term proposal to move from a survey to an administrative data basis is still largely speculative and we will not focus on this area, although we regard it as of high importance.

### ***5.2.2 The Rolling 4% Survey***

The administrative data option consists of a model initially based almost entirely on a 4% continuous rolling survey. Over time, it is expected that this survey will progressively become replaced by administrative data sources but how and when this will be done is not decided and therefore it is not possible to comment in detail on the issues involved in the process of moving between systems. Our comments will be confined mainly to that of the assumption that the source of attribute data initially will be the 4% survey.

### ***5.2.3 Quality Targets: the Rolling 4% Survey***

The 4% survey has been designed to enable estimates of appropriate precision to be obtained at local authority (LA), middle super output area (MSOA) and lower super output area (LSOA) level. The main quality target is that estimates should be available with coefficient of variation of no more than 20%. (See Appendix 2 and ONS-O2, pp. 11 and 44-45). This target is based on levels typically set both in other ONS products and other census agencies rather than based on explicit decision criteria. In many settings a 20% coefficient of variation would represent low precision. It implies that the upper bound of a 95% confidence interval is more than twice the lower bound. It is thus very important for ONS to convey to users the implications of sampling error in the survey.

As a rule of thumb, this target will provide acceptable accuracy for data on any group for which the sample includes about 32 members. If a unit of analysis contains 800 or more people in a given group, the expected number in a 4% sample would be at least 32 and therefore this group could be measured to the quality target level. If this unit contained only about 250 people, the sample of around 32 could be obtained by combining three consecutive years' data. ONS have provided illustrations of the typical accuracy with which certain types of variables at different geographical levels of the aggregation may be available.

This proposed design would involve 40% of the population being surveyed over the traditional census 10-year cycle. Even if this were to be raised to 100% with a 10% sample annually, information on population characteristics in a given year would still be subject to sampling error about only one third less than in the 4% case.

The quality target A2 (see Appendix 2) is appropriate for a continuous design such as the administrative data option and A3 for a design like the online census. The inability of the administrative data option to meet quality targets at output area level is clearly established, and

users are presented with a clear choice between timely data only down to LSOA level, or data at output area level only every 10 years.

The 4% survey is envisaged as a temporary measure in the move towards a system in which attribute data are increasingly obtained from administrative sources. If so, some attribute data could be based on close to 100% population coverage. The basis of quality targets in such a system would be very different from those which are designed for a random sample survey such as the 4% survey model.

#### **5.2.4 Content**

The most commonly required cross-tabulations will be from basic demographics – gender, age, ethnic group and relationship status. There are potential difficulties in getting reliable estimates at LA level for small ethnic groups as well as for civil partnerships and same sex married couples.

It is most unlikely that an annual 4% survey will be able to adequately cover additional addresses (i.e. address 1 year ago, second residence, workplace, student term-time address). This raises issues about the future validity of travel-to-work (TTW) area analysis even if the base units are raised from OAs to LSOAs or wards and the data is summed over a number of years. The flow matrices are so sparse and so skewed that the 4% survey could not be expected to deliver useful TTW data unless based on entirely different data sources about workplaces and travel flows.

It also puts in jeopardy reliable migration analysis, even at LA level, where as much as 10% of the population may change address annually. The NHS and DWP data proposed to be used in the creation of the SPD track migrants but present proposals do not specify whether that information may be integrated with the attribute survey. It seems unlikely that it will be possible to create workplace zones from 4% survey data for a single year. Having to combine data over too many years (say, 5) could give wrong impressions in times of rapid expansion or contraction of local employment. Such zones would have to be created from other sources such as the Inter-Departmental Business Register.

Employment/Occupation/Industry data are necessary for NS-SEC variables and it should be possible to get reasonable answers. Data about the household space such as type, tenure, rooms, bedrooms and central heating are required at all geographic levels. The rooms question has been poorly answered in past censuses and assurance needs to be made about the quality of such data. Details on access to cars and/or vans should also be adequately answered, as should major mode of travel to work, but it is possible that other administrative data could cover these issues better than the 4% survey.

Other topics may be less usefully answered by a 4% survey – even a mandatory one. The census questions of health are self-reported and although usually they are good indicators of overall health of an area they must be analysed by age/gender and hence the sample may not be great enough to capture local differences. Young carers are an important social issue but quite rare in most LSOAs therefore they may not be adequately captured.

Qualifications has always been a complicated question and would benefit from simplification. ONS needs to consult on requirements.

The census has asked a number of 'cultural' questions apart from ethnicity – country of birth, short-term migrants (STM), passport(s) held, national identity, language, English proficiency and the voluntary question on religion. There should be discussion as to whether all of these are necessary if they are only going to be determined from a 4% survey. STM may be an issue as it requires two questions based on time of arrival and expected length of stay to filter whether the individuals born overseas are usual residents or STM, who do not feature in the current definition of usual residents. If found, visitors will have to be dropped from the 4% sample.

On a practical note, there may be a need to provide internet versions of the attribute survey form in more languages than offered in 2011.

*Recommendation B18:* ONS must test which questions can reasonably be asked on a 4% internet based mandatory survey with expectation of usable answers at all geographic levels (see also Section 5.2.9).

### **5.2.5 Sample Design**

Some of these estimates will need to allow for the fact that characteristics are likely to be clustered. For example, the sample is of households rather than individuals so characteristics clustered within the household would require information on more than 32 people to meet the quality target. On the other hand, alternative sample designs may lead to increased efficiency, for example, by concentrating more of the sample in areas of high population turnover where estimates are likely to change quickly. We regard these issues as complex but relatively straightforward in that the methods available to investigate such topics are readily available. In addition, unlike in other areas, the question of optimal sample design can be treated as largely self-contained within the survey organisation itself.

Use of relevant administrative data may serve to improve accuracy, by providing information that would permit the sample design to target particular areas more efficiently. Possible sources include: change in NHS patient records, change in registered domestic addresses, new utilities customers and revised postcodes. These sources could be used as guides to LSOAs of major change in a similar way in which 2001 OAs were subdivided to produce 2011 OAs in areas of major growth. In London 550 2001 OAs had to be subdivided to get 1,550 2011 OAs. However, given the multiple uses of the survey, initially concentrating on a simple random sample design is sensible and provides a benchmark from which to assess alternative strategies. We note that objectives such as getting similar precision in different local authorities could lead to higher sampling fractions in smaller authorities.

Consideration will need to be given to integrating the designs of the 1% coverage survey and the 4% rolling survey. Sampling will need to be independent if dual system estimation is used. We would expect the coverage survey to take place over a short period of time, with the rolling survey running through the year.

We encourage ONS to continue work on sample design. Additional work should be undertaken to investigate the extent to which administrative data can be used to improve stratification and identification of areas of high population turnover to improve precision of sample estimates

especially since census data, which have historically used for such purposes will no longer be available.

### **5.2.6 Sample Frame**

Much of the infrastructure involved in conducting a 4% rolling survey will be similar to those of the online census option. Both would require the establishment of a sampling frame for both households (based on dwellings) and communal establishments. In addition, given that the content of the online survey and rolling 4% survey would be likely to be very similar, the risks appear to be both similar and well-understood for such data collection methods. Further discussion of the use of the sampling frame for developing household-based estimates is given in Section 5.2.12.

### **5.2.7 Follow-Up Design**

Whether the online census or administrative data options are chosen, it will be important to establish the accuracy of attribute data by quality checks undertaken for respondents. Both options assume an internet-led design but with phone and personal follow-up, although the precise details are not settled. A mailed paper form will be the last resort and sent out only upon request, and this option is likely to be discontinued by 2031. A 60–65% response rate from internet self-completion is assumed in line with experience elsewhere. Electronic collection will allow for quality improvements, such as sense checks in the form, the standardisation of addresses through the use of postcode lookups and possibly self-coding of questions like occupation through drop-down menus, all of which are increasingly familiar to respondents. With online response expected from a much higher fraction of the population than hitherto more attention will be needed to directly validate the impact of these set of changes, including consideration of a follow-up survey.

### **5.2.8 Estimation and Adjustment for Non-response**

Little attention seems to have been given to the possible uses of the population estimates in estimation for attributes from the rolling survey. The potential gains in precision from calibration estimation seem worthy of attention. Moreover, it is likely that there will be differential non-response even if participation is compulsory. Options for non-response adjustment using administrative data, the population estimates or other information from the SPD need to be investigated. This is a topic that has received considerable attention by the US Census Bureau with the American Community Survey.

*Recommendation B19:* Estimation methods, which make use of the population estimates or other auxiliary information from the SPD or administrative data, need to be investigated for improving the quality of attribute data, including both increased precision and for the purpose of bias adjustment for nonresponse.

### **5.2.9 Compulsory Nature of the Survey**

Analyses of both the online data and administrative data options have assumed that the census/survey would be compulsory. With response rates declining, the use of survey data without any validation such as now available from the compulsory census would be highly risky. In addition, there would be benefits in having the testing of questionnaires also compulsory. Methods for adjusting for nonresponse would be problematic in the absence of compulsory coverage data. The

arguments for this conclusion are compelling. Without the compulsory component to the survey, the risks of this method are unacceptable.

#### **5.2.10 Dissemination Policy**

There are risks in the dissemination of data from the rolling survey. Data at different geographical levels will be available in different formats, annually at the local authority level, but possibly 5 year averages in the case of LSOA data. Output area data would be unlikely to meet the quality criteria and the extent to which such data may be available is unclear. The process for release of data and advice to users about interpretation and use of data generated by this system will need care.

**Reference dates:** When data for sub-LA areas are combined after 3 or 5 years to create a sufficient sample they will be compared with LA data for the middle year of the applicable data. After three years (Y, Y+1, Y+2) it will be possible to get data for MSOAs (centred on Y+1) and after five years (Y, Y+1, Y+2, Y+3, Y+4) it will be possible to get data for LSOAs/wards (centred on Y+2). For consistency there needs to be a weighting such that the data match the basic LA demographics – gender/age/etc. – down to MSOAs (for year Y+1 as above) and hence to LSOAs/wards (for year Y+2 as above). If not there will be a plethora of upwardly summed LA estimates.

**OA data:** It is unlikely that quality OA data will be produced, even by combining ten years of sample data, but it is clear that many users at present use OA data. For some it is the finest grained building brick and is used to obtain variables for specific user-defined geographies, e.g. flood risk areas, for which LSOA level data would not be suitable. For other users the variables at OA level are themselves the basis of serious modelling – for instance the travel to work areas - or as base numbers, usually by gender and age groups, for their own service delivery or other administrative statistics.

**Recommendation B20:** ONS should consider the best way to disseminate statistics such that there is not a confusing over-supply of different statistics for an LA or MSOA relating to the same ‘central year’.

**Recommendation B21:** ONS should develop systems that can (a) prepare key OA variables (even with very high CVs) based upon five years data consistent with LSOA data and (b) enable a safe-setting whereby all variables may be combined to get user-defined geographies.

#### **5.2.11 Transition from Survey to Administrative-based Attribute Data**

Over time it is assumed that the survey will become a less substantial component and will be supplemented by administrative data based estimates. At the moment, about 150 administrative data sources are being assessed, but this work is still at a very early stage and only limited to individual-level data that is available.

One of the immediate needs is for administrative variables which could be useful for sample design and for estimation, especially to adjust for non-response bias. Here, we focus instead on the potential direct use of administrative variables as attributes.

Some variables, such as income, would provide information not previously available in the census. Some, such as qualifications, might be naturally measured from administrative sources. But there

are major concerns about how certain variables may be measured from administrative systems. For example, in census and related models, health is based on people's self-report of their status. Under the administrative data option, health may be measured in terms of receipt of benefits (ONS M-12). There are major problems in using such administrative data to measure variables such as health. For example, policies may change that move people from one form of benefit to another or are designed to reduce numbers receiving certain benefits. It is therefore difficult to separate out trends in underlying health status from those due to policy changes over time. Moreover, it would be difficult to obtain information on groups who are either ineligible for certain benefits, such as those available only to working age people, or whose health problems are such that they do not meet the threshold for receiving assistance.<sup>1</sup> If the integrity of trends is important, a period of overlap to run the alternative systems in parallel would be necessary, but if, for example, self-reported health was replaced by an administrative definition based on combining multiple data sources, this would be a very challenging process.

Some variables such as religion are not available on administrative registers and would therefore be unlikely to be available in years to come. We recommend investigation of the extent to which directly comparable attribute data for currently-available variables will be available and how the transition to a more administratively-based system might be achieved.

In order for administrative data to be used for statistical purposes, it would be necessary both for their availability to ONS be made legally binding and also that the definitions used should be such that that the information is fit for the purpose identified, such as comparability over time.

The process of transition from attribute information based on survey data to administrative data is at a very early stage and it would be necessary to undertake considerable additional work before this could be recommended. In the circumstances, it would be prudent to assume that a survey of the currently envisaged size will be retained as a major component of attribute data in the foreseeable future. There are unforeseeable risks in such a radical change and it is crucial not to implement an irreversible change until (a) there is confidence that the new system is performing sufficiently to meet well-defined quality targets and (b) there is a mechanism to monitor the performance of the new system over time to ensure that no unanticipated changes are occurring.

*Recommendation B22:* An investigation should be undertaken into the possibility of constructing one or more auxiliary variables from administrative sources, which would provide useful adjustment for nonresponse bias in the attribute survey.

*Recommendation B23:* More case studies should be undertaken, each on a single promising attribute variable which could be obtained from administrative data. The impact of quality issues, such as non-comparability across sources, on alternative analyses using the variable should be examined. For census variables, comparability with previous census estimates should be considered.

*Recommendation B24:* In any preparation work for legislation on data sharing, consideration should be given to whether the legislation can be framed in ways which would strengthen the

---

<sup>1</sup> With fully-developed register systems, use of medical and social care services, together with additional information such as use of prescription drugs can provide comprehensive estimates.

harmonization of attribute variables across administrative sources, with a view to improving their suitability for statistical purposes.

### **5.2.12 Household Estimates**

Household estimates are of high importance for many users, both as estimates and attributes. Estimates could be derived by using information on people by address from administrative sources. This would miss multi-occupancy and shared households, but if linked to survey data then there is the potential to model households. The definition of a household would need addressing if administrative data were used to derive household estimates. The social definition of a household used in the census does not apply in administrative data. Once survey and administrative sources are linked, then it will be possible to cross-tabulate administrative and survey variables.

*Recommendation B25:* More research be undertaken on obtaining consistent and comparable household information over time with the administrative data option.

### **5.2.13 Small Area Estimation for Attributes**

Research on this topic does not seem to have reached a stage where it would make any significant contribution to the principal production of attribute estimates from the rolling survey. For example, no methodology has been established for estimation for 'large numbers of variables concurrently rather than single variables on a "bespoke" basis' (ONS-M12, section 4.2). Estimation methods have also not been developed to handle the unit-level auxiliary data which would be available from the SPD. The focus has been on the kinds of small area estimation methods with which ONS is more familiar where auxiliary data is at the area level.

As is acknowledged, the value of small estimation depends upon having auxiliary variables well correlated with the attributes. We see the priority to be augmenting the SPD with administrative variables, which could be used as auxiliary variables both in calibration estimation (see section 5.2.5) as well as in small area estimation.

## **5.3 Other Outputs**

Census data have also been widely used for research purposes and for genealogical research. The online census option would permit continuation of provision of such information in essentially unchanged form, but the administrative data source would create more problems.

**LS:** The first main output is the ONS 1% Longitudinal Study (LS), which has been running since 1971. It is a valuable source of long-term information such as the primary source on trends in health inequalities based on mortality differentials. The LS is based on linkage of 1% of persons' census records, for four selected birthdates, with that of other household members, vital registration data and some other databases such as cancer registrations. To our knowledge, no work has been reported about the possibility of including the existing LS within the administrative data framework. An administrative data solution would be difficult to incorporate into the current LS design. There would be no equivalent of the decennial census data to be linked in at the same time point for all LS members. The 4% survey could potentially pick up LS members on the basis of date of birth and other information and link the attribute data. At present linkage is substantially done manually and

whether this would be possible in future is not part of our terms of reference. However, a sample survey would not include all LS members and any information would be added at random time points. We do not regard identifying and then interviewing LS members as part of the 4% rolling survey as feasible. In the longer term, if the 4% survey were to be replaced by a largely administrative data based system; this would open up a completely different set of linkages and opportunities for the LS. However, many of the variables which have been available in consistent form in successive censuses, such as self-reported health, would not be available in future. It would be possible to start a new survey based on the new system of full data linkage but this would not necessarily be in the same format as the current LS.

**SARS:** The second main type of microdata available from census for research use is the Samples of Anonymised Records (SARs). In principle, there should be no problem with providing such an extract from the internet based census or from the rolling 4% survey in the administrative data option. The latter could provide annual extracts of the rolling survey which would enhance research capability.

**100+ years data:** Census records including names are released after 100 years. These have been a valuable source of information for genealogists and historians who want to identify individuals through time. Although the information in all of the likely new systems would be paperless (apart from a few special cases), if names were attached to attribute data then comparable data to earlier periods would become available in due course. With the administrative data option, the 4% survey even if it included names of respondents would not provide the comprehensive data for genealogical studies. However, the strict rules about anonymised matching would appear to preclude names and other identifiers being linked with the data content (ONS M-9, Section 3). At present, it appears that systems that are designed to produce SPDs will not be able to produce information relevant to the needs of future genealogists and historians, although information with names could potentially be available from administrative data sources.

**Wider statistical system:** Census data also feed into many other outputs of the statistical system, for example via weights used to adjust for nonresponse in household surveys, but we have not seen it as within our terms of reference to consider such wider implications.

## 6. Transition Issues and Future Validation

A key function of the census has been to monitor change. If its methodology is to be changed significantly then it will be important for users to be able to disentangle genuine changes from artefacts of methodological change when interpreting differences between estimates from Census 2011 and estimates from the new system.

For the online census option, the two principal potential artefacts are measurement effects from the greater use of online returns and nonresponse effects from the changed data collection process, e.g. if it induces greater within-household nonresponse.

For the administrative data option, the same kinds of effects could occur with the attribute estimates, since it is anticipated that the rolling survey will make use of online data collection in a very similar way. The greater risks, however, appear to be with the population estimates under this option. The relationship between the administrative data sources and the population will change

over time in an unpredictable way with changing patterns of under- and over-coverage. As we have discussed, there are considerable challenges in adjusting for under- and over-coverage accurately. The adjustment methodology, based upon the relation between the experimental SPDs and Census 2011, may no longer be appropriate as time advances beyond 2011.

It therefore seems essential that methods are planned for checking the validity of future estimates, under either option. Such checks are needed not only for users, but also to inform methodological development.

In Recommendation B3, we have already encouraged ONS to investigate quality issues associated with the use of internet data collection. These issues include mode effects associated with both measurement and changed patterns of nonresponse. Empirical research into such mode effects would be valuable for either option and we encourage it to be planned with some urgency. In addition to the need for research to inform the implementation of the option, we also recommend planning future validation checking to investigate the nature of any methodological artefacts of change which arise on implementation.

We now turn our focus to future validation of population estimation under the administrative data option. Whereas the Programme has been able to compare estimates from experimental SPDs with those from Census 2011, this option would not be available were the census to be discontinued. It seems essential for mechanisms to be in place to detect problems with the estimates. Given the radical change in population estimation methodology under this option, we feel that it is not only necessary to be able to detect problems but also to have a suitable methodology in place to address problems if they do arise.

An established approach when introducing major methodological changes in statistical series is to run both systems in parallel for a period of time. The resulting parallel sets of estimates help users to understand the effect of the methodological change and even permit some degree of calibration, where users can adjust the new estimates to the old methodology or vice versa. To a limited extent the mid-year estimates based upon Census 2011 could serve that purpose but their value will run out. The most direct way to run options in parallel would be to run both the online census option and the new administrative data option in 2021, although the feasibility of that seems to be much reduced by not only cost considerations but also from the availability of ONS staff to run two major exercises.

If there is not to be a census in 2021, we consider it as essential that there be a substantial validation exercise, developing the idea proposed of a 10% coverage survey in 2021 (ONS-O2, section 5.1), as well as further exercises in the future. This would be in addition to the kinds of checks described in section 5.1.9. We think the validation exercise should be set up in such a way that, not only can it detect unsatisfactory population estimates from the administrative data option, but also it provides an alternative basis for estimation if this occurs. This is recommendation A4 below. Ideally, this exercise would provide users with some idea of the impact of methodological differences (between Census 2011 and the new approach) on the population estimates in 2021 and even enable some degree of calibration across the methodologies. This might be achieved by using some version of Census 2011 methodology in the coverage survey. It does not seem essential that this exercise is tied to the single year 2021 - it could even be spread over more than one year and would need to be planned into the future.

In summary, given the tremendous importance of using estimates of the kind produced from the census for comparisons over time, we believe that an essential step, if the administrative data option is to replace the census, is to invest in some form of substantial ongoing validation exercise.

*Recommendation A4.* ONS should establish a methodology for checking the validity of the population estimates and for producing alternative estimates when these estimates are deemed unacceptable. We expect that this methodology will require substantial additional data collection.

## 7. Conclusions

A key question addressed by the report is whether each of the two front running options provides a methodologically sound basis for replacing Census 2011 methodology, given the quality standards specified by ONS.

Our answer to this question is positive for the **online census option**, subject to two provisos. First, we assume that all cases where online responses are not received will be followed up fully, using a similar level of investment in follow-up data collection to Census 2011. Second, we assume that thorough research will be undertaken into possible mode effects (measurement and nonresponse) of the interventions adopted to achieve greater internet response and that lessons from this research will feed into the development of the methodology for this option (see Recommendation B3).

Regarding the **administrative data option**, we do not believe that the Programme has yet reached a stage where we can answer the above key question positively. The Programme has undertaken some valuable research about the option and has taken a number of steps to demonstrating its feasibility. In particular, it has shown that administrative sources may be linked at the individual level even after the data are anonymised and it has demonstrated that a dataset could be constructed from administrative data sources in 2011 where the population counts in local authorities in the dataset show a close relation to the population estimates from Census 2011. These results and other research show that the option has some promise.

Before there is sufficient evidence to answer the key question for the administrative data option, we consider that the work proposed in Recommendations A1, A2, A3 and A4 (see earlier in report) needs to be undertaken.

In addition to obtaining satisfactory conclusions to the matters identified in these four recommendations, we consider that two items of legislation will be required for the answer to the key question to be positive.

First, there will need to be data sharing legislation which ensures that the administrative data sources underlying this option will meet the needs of statistical estimation. The aim is to ensure that changes in population estimates should not result from changes in the administrative systems rather than from genuine changes in the population. The National Statistician will need to be given influence over the content and management of the administrative systems.

Second, the rolling survey and Population Coverage Survey will need to be mandatory. Response rates in sample surveys have been declining relentlessly and bias from nonresponse is a major risk when producing estimates from surveys. We view it as essential for these surveys to be mandatory if the quality of the resulting estimates is to meet ONS's standards.

Recommendations A1-A4 relate to the immediate decision about whether to proceed with the option. If it is carried forward there will be further decision points which will require further evidence. In particular, the Programme has not yet provided evidence as to whether the introduction of administrative data or small area estimates for attributes could enable the scale of the 4% rolling survey to be reduced, with consequent savings, without overall loss of data quality.

We need to stress again that our recommendations relate only to the methodological basis of the programme, not to its costs nor to its benefits for users. From a user perspective, a key difference between the administrative data option and the online census option is that the sampling employed in the rolling annual survey of the former will prevent the accurate estimation of many very small area or small population characteristics which would be feasible from the latter option. On the other hand, estimates would be available from the administrative data option annually, compared to only once every 10 years for the census option. Which of the two options a user prefers is likely to depend on the relative importance attached to these two benefits, i.e. accurate small area population estimates and characteristics every 10 years vs. less detailed estimates annually.

Although we consider that the key work required to decide about the potential methodological soundness of the administrative data option is contained within recommendations A1-A4, we also recommend considerable additional work, listed under recommendations B1-B25. Some of these recommendations deal with fundamental questions about the methodology of the options. In general, they are intended to guide prioritisation of further work.

Finally, we should like to commend ONS for some excellent work on the Programme and for carrying through their identification and evaluation of options in such a well-organised way. Our recommendations for further work on the Programme overlap quite considerably with proposals for next steps within the Programme documents which we have read and we consider that ONS has a strong understanding of the methodological challenges and risks it faces.

## About the Authors

Chris Skinner is Professor of Statistics at the London School of Economics. He is a Fellow of the British Academy and of the American Statistical Association and was awarded a CBE in 2010. He was a member of the UK Census Design and Methodology Advisory Committee (2005-10), the National Statistics Methodology Advisory Committee (2001-10) and the Statistical Methods Advisory Committee of Statistics Canada (2000-11).

John Hollis is a demographic consultant formerly at the Greater London Authority. He was the President of the British Society for Population Studies (2005-07). He led the local government representatives on the CLIP Census Advisory Group for the 2001 and 2011 Censuses and was a member of the 2011 Census High Level Quality Assurance Panel.

Mike Murphy is Professor of Demography at the London School of Economics. He is a Fellow of the British Academy, Research Secretary of the Population Investigation Committee and past-President of the British Society for Population Studies (2003-5). He was a member of ONS Working Group on Content and Steering Committee for 2001 One-Number Census, and Steering Group Review of Civil Registration Statistics.

## Acknowledgements

We are very grateful to James Brown, Tony Champion, Ian Diamond, Piers Elias, Tim Holt, David Martin, Les Mayhew, Philip Rees, Natalie Shlomo and Ludi Simpson for comments on a draft version of this report. All views expressed in the report are, however, those of the three authors alone.

## References

### ***Beyond 2011 Reports and Publications***

(available on [www.ons.gov.uk/ons/about-ons/who-ons-are/programmes-and-projects/beyond-2011/reports-and-publications/index.html](http://www.ons.gov.uk/ons/about-ons/who-ons-are/programmes-and-projects/beyond-2011/reports-and-publications/index.html))

ONS-C2 Summary of the Uses of Census Information, September 2013

ONS-M1 Options Explained, October 2012

ONS-M2 Exploring the Challenges of Using Administrative Data, 2012

ONS-M3 Options Explained 2, February 2013

ONS-M5 Options Explained 3, July 2013

ONS-M7 Producing Population Estimates using Administrative Data: In Practice

ONS-M8 Producing Population Estimates using Administrative Data: In Theory

ONS-M9 Matching Anonymous Data, July 2013

ONS-M12 Producing Socio-Demographic Statistics, September 2013

ONS-O1 Options Report, February 2013

ONS-O2 Options Report 2, July 2013

ONS-P2 Progress Report, February 2013

ONS-P3 Progress Report, July 2013

ONS-R2a Exploring the Feasibility of a Rolling Census Design for England, J. Brown, February 2013

ONS-R3 Administrative Data Models – Research Using Aggregate Data, M. Ralphs, M. Elkin, S. Whitworth, J. Wall, D. Rasulo and J. Huynh

ONS-S1 Administrative Data Sources Report: NHS Patient Register, November 2012  
ONS-S2 Administrative Data Sources Report: Electoral Register, February 2013  
ONS-S3 Administrative Data Sources Report: The English School Census and the Welsh School Census, February 2013  
ONS-S4 Administrative Data Sources Report: Higher Education Statistics Agency: Student Record, May 2013  
ONS-S5 Administrative Data Sources Report: CIS and L2, July 2013

### **Other References**

Administrative Data Taskforce (2012) The UK Administrative Data Research Network: Improving Access for Research and Policy. Report published by ESRC, MRC and Wellcome Trust

Ben Hur, D. and Burck, L. (2013) Rolling Integrated Census in Israel. Paper presented at World Statistical Congress, Hong Kong.

Brown, J.J., Diamond, I.D., Chambers, R.L. Buckner, L.J. and Teague, A. D. (1999) A methodological strategy for a one-number census in the UK. *Journal of the Royal Statistical Society, Series A*, 162, 247-267.

Christen, P. (2012) *Data Matching: Concepts and Techniques for Record Linkage, Entity Resolution, and Duplicate Detection*. New York: Springer.

Economic Commission for Europe. (2012) Conference of European Statisticians Sixtieth plenary session, Paris, 6-8 June 2012 Overview of the 2010 round of population and housing censuses in the United Nations Economic Commission for Europe region. Note by the secretariat. United Nations ECE/CES/2012/42/Rev.1.

Harper, G. and Mayhew, L. (2012) Using administrative data to count local populations. *Applied Spatial Analysis and Policy*, 5, 97-122.

Hogan, H. (1993) The 1990 Post-Enumeration Survey: Operation and Results. *Journal of the American Statistical Association*, 88, 1047-1060.

Hogan, H. (2003) The accuracy and coverage evaluation: theory and design. *Survey Methodology*, 29, 129-138.

Jeskanen-Sundström, H. (2008) Overview and challenges in the use of administrative data in official statistics. Paper presented at conference of International Association of Official Statistics, Shanghai.

Large, A., Brown, J., Abbott, O. and Taylor, A. (2011) Estimating and Correcting for Over-count in the 2011 Census. *Survey Methodology Bulletin*, 69, 35-48.

Mulry, M.H. (2007) Summary of Accuracy and Coverage Evaluation for the U.S. Census 2000. *Journal of Official Statistics*, 23, 345-370.

Nirel, R. and Glickman, H. (2009) Sample surveys and censuses. Ch. 21 in Pfeffermann, D. and Rao, C.R. eds. *Sample Surveys: Design, Methods and Applications*. Handbook of Statistics, Vol 29A, Elsevier.

ONS (2004) 2001 Census: Manchester and Westminster Matching Studies Final Report. September 2004

ONS (2012a) Census update February 2012: providing the online census

ONS (2012b) Response Rates in the 2011 Census. 2011 Census: Methods and Quality Report, December 2012

ONS (2013) Labour Force Survey: Performance and Quality Monitoring Report. April to June 2013.

- Ralphs, M. and Tutton, P. (2011) International models for census taking: current processes and future developments. Office for National Statistics (available on Beyond 2011 website)
- Renaud, A. (2007) Estimation of the coverage of the 2000 census of population in Switzerland: Methods and results. *Survey Methodology*, 33, 199-210.
- Statistics Canada (2012) Internet data collection in the Canadian census of population. Paper presented at 14<sup>th</sup> Group of Experts on Population and Housing Censuses, Conference of European Statisticians, UN Economic Commission for Europe, Geneva.
- Teague, A. and Swan, J. (2011) Beyond 2011 – list of statistical options. Office for National Statistics (available on Beyond 2011 website)

## **Appendix 1. Terms of Reference**

### **Introduction**

The Office for National Statistics (ONS) is currently taking a fresh look at options for the production of population and small area socio-demographic statistics for England and Wales. The Beyond 2011 Programme has been established to carry out research on the options and to recommend the best way forward to meet future user needs.

Improvements in technology and administrative data sources offer opportunities to either modernise the existing census process, or to develop an alternative by re-using existing data already held within government. Since methods for taking the traditional census are already relatively well understood most of the research is focussing on how surveys can be supplemented by better re-use of 'administrative' data already collected from the public.

The final recommendation, which will be made in 2014, will balance user needs, cost, benefit, statistical quality, and the public acceptability of all of the options. The results will have implications for all population-based statistics in England and Wales and, potentially, for the statistical system as a whole. In order to work towards the deadline for the final recommendation in 2014 an emerging view will be taken in July 2013. This emerging view will be based upon a body of evidence to develop, investigate and evaluate a range of different options.

A Quality Assurance (QA) process for the methodological aspects of the Beyond 2011 Programme has been established. The aim of this QA process is to ensure that the statistical work and research conclusions that make up this body of evidence to support the recommendations are independently and rigorously reviewed. This process is essential to ensure that the methods developed within the Beyond 2011 Programme are statistically robust. One aspect of this QA process is this final review panel. This will provide ONS with a valuable opportunity to take stock of research to date and the evaluation of the options and hence recommendations made and to identify any areas where evidence needs to be strengthened.

### **Objectives**

The objectives of the review are to:

- provide independent assurance and thus confidence to users that the methodology and statistical research behind the recommendation in 2014 is based upon sound research and a strong evidence base;
- identify areas where evidence could be significantly improved;

- identify the main risks with the emerging view, noting any particular problems/challenges with the methods recommended, and any further work required to mitigate these risks and challenges.

## Scope

The 8 initial Beyond 2011 options for the production of population and small area socio-demographic statistics are:

1. Full census
2. Rolling census
3. Short form plus long form
4. Short form plus annual survey
5. Administrative data-based models - aggregate level
6. Administrative data-based models - individual level with partial linkage
7. Administrative data-based models - individual level with full linkage
8. Address register plus survey

It will be within scope to review:

- the initial evaluation of each of these options and the journey from these 8 to a short list of 6 options
- the methodological detail of how each of these 6 options could be implemented
- detailed evaluation of the 6 options, specifically focusing on how each option meets quality criteria for both population and small area socio-demographic statistics
- the evaluation criteria and quality criteria themselves

Documentation on all of these areas will be available to ensure that the reviewers have a complete understanding of the research. The review will focus on the development and evaluation of options by ONS for England and Wales.

It will be out of scope to review:

- detailed operational issues related to the different options, other than those that relate to specific evaluation criteria
- work undertaken within the stakeholder and engagement project of the programme, including research on public acceptability
- work undertaken on the future production of population and small area socio-demographic statistics in Scotland and Northern Ireland

## Reviewers

It is expected that no single person will have the expertise or experience required to conduct the review alone. Therefore it is expected that a small team consisting of two or three experts will conduct the review.

Ideally the review team will need the following experience and expertise:

- Survey design in the context of population estimation
- Use/analysis of administrative data
- Statistical modelling/record linkage
- Demography and specifically population estimation
- Experience in the use of population and small area socio-demographic statistics

## Timing

The entire review should be completed by end of September 2013, with a report delivered to ONS. Publication of the report will follow in early October. An indicative timetable for the review is detailed below.

June (pre-review information gathering)

- Understand background and motivation to the Beyond 2011 Programme
- Review initial evaluation of 8 options
- Review initial research and first evaluation of 6 options (as published in February)

July - September (formal review period)

- Detailed evaluation of research, evidence and recommendation
- Presentation of emerging findings (mid Aug)
- Produce a report summarising the findings of the review (end of September)

## Liaison with ONS

The review team will need to have frequent access to ONS to help understand particular aspects of the work, request further information and report on progress. The key contacts at ONS will be:

- Peter Benton - Beyond 2011 Programme Director - overall responsibility for the review
- Jane Naylor - Principal Methodologist - manage contract, liaison with the review team
- Andy Teague - Head of Statistical Options Development Project - supply/exchange of information

## Appendix 2. ONS Quality Standards (Source ONS-O2)

### Local authority quality standards for population estimates

	<b>97 per cent of local authority population estimates have a 95 per cent confidence interval of ...</b>	<b>All local authority population estimates have a 95 per cent confidence interval of ...</b>
<b>P1</b>	+/- 3.0 per cent or better	+/- 3.8 per cent or better
<b>P2</b>	+/- 3.0 per cent or better in the peak year +/- 6.0 per cent or better in year nine	+/- 3.8 per cent or better in the peak year +/- 13.0 per cent or better in year nine
<b>P3</b>	+/- 5.2 per cent or better	+/- 8.5 per cent or better

### National quality standards for population estimates

	England and Wales population estimates have a 95 per cent confidence interval of ...	England population estimates have a 95 per cent confidence interval of ...	Wales population estimates have a 95 per cent confidence interval of ...
<b>P1</b>	+/- 0.15%	+/- 0.15%	+/- 0.64%
<b>P2</b>	+/- 0.15% in the peak year +/- 0.27% in year nine	+/- 0.15% in the peak year +/- 0.28% in year nine	+/- 0.64% in the peak year +/- 1.06% in year nine
<b>P3</b>	+/- 0.22%	+/- 0.23%	+/- 0.81%

### Quality standards for population attributes

	Geography and frequency	Quality standard (accuracy)
<b>A1</b>	LA and MSOA data once every 10 years “Do minimum”	LA statistics for an attribute applying to 0.2% of the total LA population have a CV of 20% or less MSOA statistics for an attribute applying to 3% of the total MSOA population have a CV of 20% or less
<b>A2</b>	LA data annually based on 2 year rolling average MSOA data annually based on a 3 year rolling average LSOA data annually based on a 5 year rolling average “Greater frequency”	Exceeding the accuracy targets for A1 plus: LSOA statistics for an attribute applying to 10% of the total LSOA population have a CV of 20% or less
<b>A3</b>	LA, MSOA, LSOA and OA data once every ten years “Greater detail”	Exceeding the accuracy targets for A1 plus: OA statistics for an attribute applying to 10% of the total OA population have a CV of 20% or less