

**Economic and Social Council**Distr.: General
5 May 2010

Original: English

Economic Commission for Europe

Conference of European Statisticians

Group of Experts on Population and Housing Censuses**Thirteenth Meeting**

Geneva, 7-9 July 2010

Item 4 of the provisional agenda

Beyond the 2010 census round: Plans for the 2020 round**Beyond 2011: Building a population statistics system for the future****Note by the Office for National Statistics, United Kingdom***Summary*

In recent years substantial work on improving United Kingdom population statistics has taken place, with a strong focus on the use of administrative data to deliver improvements in the mid-year estimates of the population. At the same time significant improvements are planned for the 2011 Census, especially in relation to migration statistics.

Although these planned improvements will deliver better the statistics which better meet user needs, the National Statistician, together with the Registrars General, recognised the need for a longer-term programme of work which would build on current improvements. To meet this need they established the Beyond 2011 project which would look at options for the population statistics system in the long term.

The Beyond 2011 project has started feasibility work on a number of options which could provide alternative or complementary sources of population and core socio-demographic statistics. The focus of the work is on integrating administrative data, extended survey approaches and alternative methods of Census taking, and looks at how these different data sources together could best meet the need of statistics users. This paper outlines the background to Beyond 2011 and the early work around options and models.

I. Introduction

1. Basic statistics on the population are one of the key outputs for any National Statistics Office. In the United Kingdom the Office for National Statistics, and its predecessors such as the OPCS (Office for Population, Censuses and Surveys) has, throughout the years, provided the key statistics on the population, its structure and its characteristics.

2. Since 1801 the decennial Census of population has delivered a comprehensive snapshot of the population every ten years. To complement the census and to provide up-to-date information in between Censuses mid year population estimates are produced each year, based on the last census. They are then rebased following the next census.

3. Although this approach has generally met the needs of a wide range of users it is clear users increasingly want outputs at low levels of geography to be available more frequently to have a more accurate picture of population change – how do people and households change over time, where do they move to and from, and which characteristics do they have? To meet this need for more detailed up-to-date information substantial recent cross-governmental efforts have led to improvements in inter-censal estimates of the population and its characteristics (especially with respect to migration statistics). Nevertheless it is clear these improvements need a coherent framework of sources to ensure cost-effective, coherent and sustained improvements in the long term.

4. The Beyond 2011 project sets out the initial work which addresses the key difficulties in measuring an increasingly mobile and changing population by focusing on the integration of existing and new data sources. It will investigate options for the production of statistics in a new analytical environment which will need to meet existing and new user needs in policy making and the wider user community.

5. This paper sets out the current way in which our population statistics are produced, and highlights some of the key concerns users currently have in relation to their fitness for purpose. We then outline the Beyond 2011 work programme to take place over the next few years and we will give an overview of potential data sources, data models and their advantages and disadvantages.

II. Current population statistics, population change and changing user needs

A. Current data on the population

6. The Office for National Statistics produces a wide variety of data on the population and its characteristics. Although the decennial Census is a key data source for a wide range of statistical outputs there are other data sources such as surveys and administrative sources which are routinely used to produce population data.

7. Currently statistics are produced within various broad domains:

- (a) Population counts (e.g. counts by geography, sex and age)
- (b) Population flows (births, deaths and gross migration flows)
- (c) Population structures (e.g. family and household relationships)
- (d) Population and housing characteristics (e.g. health, employment)

8. The decennial Census provides a comprehensive population count at very fine levels of geography, together with information on population structures and characteristics.

9. In the years in between the Census the ONS Centre for Demography (ONSCD) produces annual mid-year estimates of the population through the established cohort component methodology, using a combination of survey data and administrative data sources.

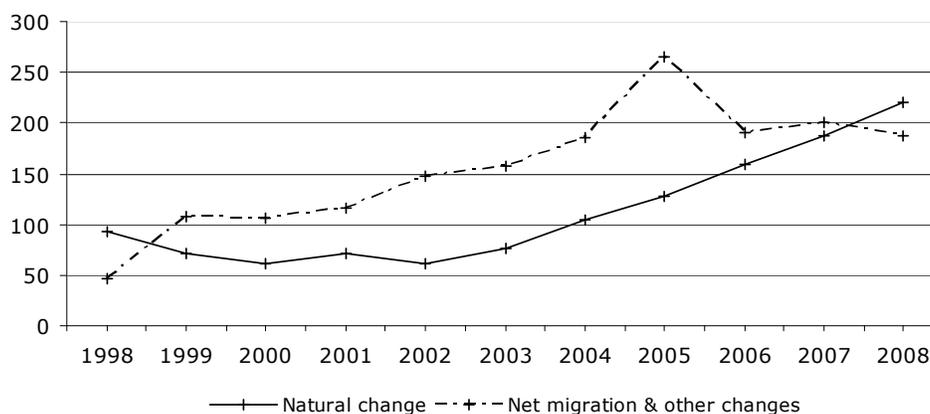
B. Population change and changing user needs

10. The rate of change in the population of the United Kingdom has been increasing over the last few decades. The greater mobility of the population has meant that more people move in and out of the country than ever before. At the same time, there have been transformations in the other components of population change. Both birth and death rates have declined over the last ten years, and international migration has taken over as one of the key drivers for United Kingdom population growth, although this trend changed again in 2008. These changes in the components of the United Kingdom population are shown in Figure 1 below.

Figure 1.

Components of population change, mid-1998 to mid-2008, United Kingdom

Thousands



11. These rapid changes make it increasingly challenging for statisticians to accurately count the population on a frequent basis and at low levels of geography.

12. Traditionally, in most countries a census has been the key statistical source to provide a population count. Its advantages are well rehearsed: at a given point in time and with regular frequency (usually every ten years) it provides a rich and comprehensive picture of the whole population in a country. However, for many countries it has become clear that a census count alone cannot provide the information on the population which policy makers and other users need.

13. In addition there has been an increased recognition of the difficulties of this traditional approach in measuring population change and meeting user needs¹. One of the

¹ For a comprehensive overview of the issues, see House of Commons Treasury Committee, Counting The Population, Eleventh Report of Session 2007-08. London: The Stationery Office.

compounding problems is the decreasing willingness of the public to cooperate with data collection exercises such as household surveys and censuses, as reflected in lower response rates². At the same time it is clear administrative data sources are increasingly becoming available within government. Although these data are collected for administrative rather than statistical purposes, there is some recognition of their potential to provide basic statistical information.

14. To respond to this need for need for more frequent and timely information and the drive for public sector efficiency, some countries have developed statistical counts of the population through means other than a census. Scandinavian countries in particular, have rich and well-established register-based systems that are better suited to counting the population more frequently and efficiently. These administrative data systems, as the name implies, are widely used for public service delivery, but they are also used for statistical purposes. Other countries, such as the United States America and France, have moved away from a traditional long form Census and provide population statistics through a combination of a short form and survey approaches (USA), or through a rolling Census (France).³

15. As discussed earlier the United Kingdom does not have a register-based system, and uses a census approach to count the population every ten years. In addition, a range of surveys and administrative data are used to update the population count in between censuses at each mid-year point through the population mid-year estimates. However, it is clear the fitness for purpose of this approach is being questioned, and the combined challenges of a changing population, changing user needs, and changes in data availability point to the need for an open debate with users on their requirements and priorities, and how they can best be met.

16. In meeting these challenges the United Kingdom National Statistician, together with the Registrars General in the devolved administrations, set up the Beyond 2011 Project. This project, led by the Office for National Statistics but supported by the wider Government Statistical Service, aims to assess the feasibility of improving United Kingdom population and other social statistics through alternative data sources which could complement or replace existing approaches.

III. Beyond 2011: responding to the need for change

A. Aims and objectives

17. The high-level aims of the Beyond 2011 work programme are:

(a) to investigate the feasibility of improving population statistics in the United Kingdom by making use of integrated data sources to replace or complement existing approaches;

(b) to investigate whether alternative data sources can provide the priority statistics on the characteristics of small populations, typically provided by a Census.

² For example, see Martin,J, and Matheson,J. (1999) "Responses to Declining Response Rates on Government Surveys", Survey Methodology Bulletin, 45 (1999). London: Office for National Statistics

³ For an overview of different approaches please see <http://unstats.un.org/unsd/demographic/sources/census/alternativeCensusDesigns.htm>

18. The central focus of the work will be to consider how to develop population data of the required coverage and quality as efficiently and effectively as possible. The project will need to establish which data are needed to support the policy, business and administration needs of the many key users, not least public sector users. Given the known budgetary pressures and public concerns, there will need to be a clear emphasis on how best to collect the minimum necessary data about the population at the minimum cost whilst meeting public concerns about security, privacy and intrusiveness.

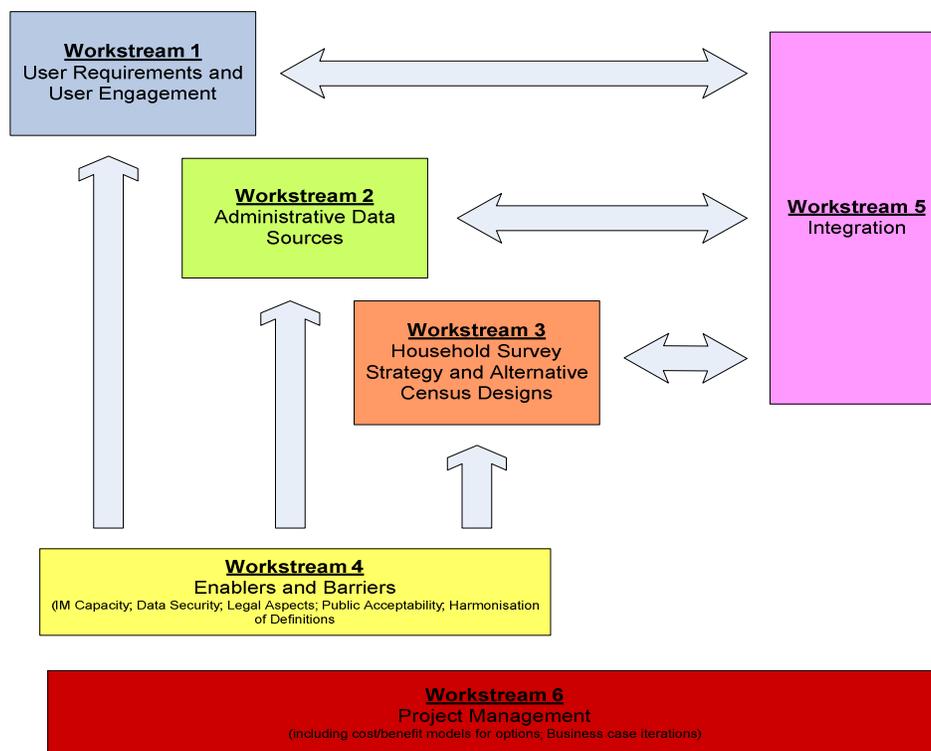
B. Beyond 2011 work programme

19. In developing the work programme for Beyond 2011 a number of key issues, both statistical and non-statistical, will need to be addressed. The work will need to focus on the technical/statistical development of alternative options but also on stakeholder engagement, user requirements, public engagement and operational delivery. In addition significant work will need to take place on cost / benefit analysis of alternative options to ensure solutions can be judged on their efficiency and user acceptability compared to the current system.

20. The work for Beyond 2011 has been structured into six work streams as shown in Figure 2.

Figure 2.

Beyond 2011 Work Streams



21. Workstream 1 will engage with social statistics users in order to establish their current and likely future requirements for population, migration and other statistics. This will cover not just which statistics are required, but also the quality dimensions such as accuracy, levels of geography, frequency and timeliness with which they are required.

22. Workstream 2 will identify the existing administrative data sources ONS will need access to. This is in order to deliver a feasibility assessment on how administrative data could be used to create a population count.

23. Workstream 3 will investigate other approaches, over and above the use of administrative data, that may also be needed. These are:

- (a) The contribution that can be made by household and individual surveys;
- (b) Alternative Census designs (including looking at approaches used by other countries);
- (c) Synthetic / small area estimation.

24. Workstream 4 will address a number of issues that may act as enablers or barriers to the three previous workstreams. These are as follows:

- (a) IM capacity requirements
- (b) Data security
- (c) Legal aspects of bringing together various types of data sources
- (d) Public acceptability of data linking within government
- (e) Harmonisation and conceptual issues across sources

25. Workstream 5 will bring together the findings from workstreams 1-4 to deliver a detailed scoping study on options for a population statistics system to produce population, migration and other statistics typically provided by a Census which meet user needs.

26. Workstream 6 will focus on the development of cost/benefit models for options, in parallel with the statistical and technical development work in the other workstreams and the work on user requirements. This workstream will also comprise the related work on further business case iterations and the overall project and work programme management.

C. Key issues and risks

27. In the development of the Beyond 2011 work it is clear a number of risks and issues will need to be actively managed.

(a) Although the decennial census has shortcomings it is a widely used source for population and socio-demographic statistics. As with previous censuses the first part of the work programme for the 2011 Census was a review of the need for a census⁴. This concluded that currently there is a continuing need from all users for information which can at present only be produced by a conventional census. There is therefore a risk that public communication on alternatives to a census or modifications to the current approach has an adverse effect on the success of the 2011 Census which takes place on 27th March 2011. User engagement tasks will therefore need to be very carefully planned and carried out in order to minimise this risk.

(b) The work programme will propose to bring together a number of data sources. However, there currently is no address or population register that could be used as a 'spine' for a population count. Other countries which have moved away from a census approach have robust registers in place for housing stock (addresses) businesses and the population. The lack of comprehensive registers in the United Kingdom significantly

⁴ <http://www.ons.gov.uk/census/2011-census/2011-census-project/case-for-2011-census/index.html>

increases the risk that some people will be omitted from an alternative statistical system, and also it makes integrating multiple data sources much more costly and subject to error.

(c) The successful development of the feasibility work leading to an integrated system of data sources will be dependent on intensive and sustained cross-governmental efforts in data sharing and co-ordination of policy needs. Although the Statistics and Registration Service Act 2007 offers new opportunities to set up legal gateways for data sharing for statistical purposes, the integration of data from different sources remains a challenging task.

(d) The statistical outputs from a new integrated system are likely to be of a different quality from current sources such as the census. For example, it is likely some outputs will be available more frequently but in less detail. User engagement will be key to determine which properties are more important for which outputs but it is inevitable there will be a trade-off between the different quality dimensions. This will need careful consultation.

(e) The delivery of this work will be complex in both its statistical and operational aspects. Apart from a sustained cross-governmental effort complex technical issues will need to be overcome – for example in information system support and data security.

(f) Although it is not the aim of Beyond 2011 to build and maintain a population register in its own right the efforts to create a comprehensive data source for statistical purposes may be perceived as excessively intrusive. Public and political acceptability of the proposals will need to be assessed and taken into account.

IV. Alternative data sources, options, and models

A. Data sources

28. As discussed before the basic aim of Beyond 2011 is to determine the extent to which a new approach might form an alternative basis for producing population and core socio-demographic statistics which are fit for purpose. At a conceptual level there are three different components to a possible new approach:

- (a) Census data
- (b) Survey data
- (c) Administrative data

29. It is clear each of these data sources have their strengths and weaknesses. The following sections give a high-level overview of the main issues.

1. Census data

30. The advantages of a Census are well rehearsed: a census provides a rich statistical data source, with key information on individuals and households at the unit-level. This means not only univariate analysis but also multi-variate analysis can take place for very low levels of geography (but taking into account respondent confidentiality). It usually provides good accuracy, and different methods exist to check and quality assure the actual census count (either through dual-system estimation and/or adjustments from administrative data sources). In addition a traditional Census approach is familiar to both users and producers of population and wider socio-demographic data: the strengths and weaknesses are well known and are usually taken into account in their use.

31. On the other hand, it is clear the traditional Census approach has a number of disadvantages. It only provides a single snapshot of the population, at one point in time, and with a significant time gap (usually five or ten years). It is generally perceived as expensive and burdensome as information needs to be provided by individuals and households – indeed it is clear the willingness of the general public to cooperate on censuses and surveys is declining and extra resources need to be spent to achieve high levels of response.

2. Survey data

32. One of the key advantages of household surveys is that they are able to collect very rich information about the various characteristics of the population, and as such they can be efficient in the amount of information which can be collected for a relatively low cost. They can ask information on complex characteristics, and can be relatively reactive in meeting new and emerging information needs. Surveys can also be run on a continuous basis and provide good data on trends and changes, especially in the case of longitudinal surveys.

33. The key disadvantage of a survey approach is the lack of information on the full population: information is only available for a sample of the population, so information is not available at very small levels of geography. This also means survey outputs are estimates rather than counts, and users need to take account of measures of uncertainty through estimates of sampling error. It is also clear there is some burden on households and individuals to provide survey responses, especially in the case of long and/or complex survey questionnaires. In the last few decades it has become clear the public is less willing to co-operate with surveys, which had led to lower response rates, increasing the risk of non-response bias.

3. Administrative data

34. Administrative data are usually collected for administrative, rather than statistical purposes. However, when they are used for statistical purposes it is clear they are relatively cheap as their collection follows from existing processes and / or systems. This also means they are less burdensome, and often are available on an ongoing basis. Because of their nature the statistics they can provide are very often available at low levels of geographical granularity. In addition there have been considerable advances in the scope, completeness and quality of administrative sources over the last decade and there has been on-going work on improving reconciliation between administrative sources through pilots of data linkage⁵.

35. On the other hand it is clear that because they are not collected for statistical purposes very often extra processing and / or estimation needs to take place to make them fit for producing statistical outputs. For example, often lots of work needs to take place on matching and linking unit-level records, and often duplicates need to be removed. Over and above basic details they are not usually able to provide detailed information on person characteristics, and when they do definitions and / or coverage are often different compared to other sources or to the population of interest. Usually statistical producers do not have control over the data source which leads to uncertainty over their ability to provide ongoing time series. There are typically numerous issues in relation to the legality of their use for anything other than administrative purposes, the sharing between owners and other ownership issues, and (perceived) security and privacy issues.

⁵ See, for example: <http://www.ons.gov.uk/about-statistics/methodology-and-quality/imps/updates-reports/current-updates-reports/feasibility-linkage-of-births-records-to-school-census-records---march-2010.pdf>

B. Data models and potential options

36. From the brief description of each data source's strengths and weaknesses it is clear any alternative solution to the current system of population statistics production will need to rely on a combination of data sources to provide a way forward. However, one of the challenges in the Beyond 2011 work will be to establish exactly which combination would best meet user needs for accurate and relevant population statistics in the most efficient and effective way.

37. In the early thinking on potential options for the future, work has taken place on alternative models in other countries. The following broad categories of options relate to the ideal, eventual solution for producing such statistics. There are separate issues around how one might move from the current solution to this future solution. It might be possible to do this in a single step or several interim steps might be required.

1. Category 1 – A point in time Census provides a population spine

38. Under this category a point in time Census would still be conducted to provide at least the basic population spine and possibly additional information. This might need to be supplemented by surveys and/or administrative data sources to 'fill the gaps' in the information available from the Census (e.g. to provide the topic estimates to supplement a short-form census or improve the accuracy of estimates of long form variables).

2. Category 2 – Administrative data is used as the basis for population estimates

39. Under this approach administrative data would replace a traditional Census in providing the basis for producing population estimates, either through creating an administrative data based population spine (unit level administrative records) or using low level aggregates to provide population counts. Administrative data could also provide some topic information. Estimates that cannot be produced from administrative data sources would be produced based on a survey approach. As the administrative data sources would have been compiled at different times, this approach moves away from the point in time approach of a traditional Census.

3. Category 3 – A survey or rolling census approach is used

40. This approach also moves away from the point in time approach of a traditional Census to producing estimates on a continuous basis. These estimates could be based on either a rolling Census or survey design, supplemented by administrative data where this could improve the accuracy of the estimates produced (either as a covariate in estimation or for quality assurance purposes).

V. Conclusion

41. Although the work on Beyond 2011 is in its early stages it is clear there is a strong need to thoroughly develop a work programme which develops and assesses a number of alternative options to the current model of population data sources. In the early thinking on options the project has taken into account alternative models in use in other countries, and further work is ongoing to develop the detailed options. At this point in time however it is clear a wide range of options will need to be developed and assessed further, both on its statistical quality and fit with user needs, and on their relative costs against their benefits.

42. It should be stressed the work on alternative options does not imply a decision on a traditional 2021 Census. It is not expected the work on options development will conclude

until 2014, at which point a decision on a 2021 Census would need to take into account the evidence on feasibility, or otherwise, of alternative approaches.
