

Discussion paper

Proposals for an Integrated Population Statistics System

Office for National Statistics, October 2003
© Crown Copyright

Contents

	Page
1 Background	3
2 An Integrated Population Statistics System	3
3 Benefits of the population statistics system	4
3.1 Statistical benefits	4
3.1.1 More frequent and timely small area population statistics	4
3.1.2 More consistent population statistics	4
3.1.3 More detailed population statistics	4
3.2 Benefits to society	4
3.2.1 Improved policy formulation and monitoring	4
3.2.2 Improved service delivery	4
3.2.3 Improved efficiency	5
3.2.4 Reduced public burden	5
4 Risks to the proposed Integrated Population Statistics System	5
4.1 The legality of linkage	5
4.2 Political and public acceptability	5
4.3 Quality and availability of administrative data	5
4.4 Dependencies	5
4.5 Confidentiality and disclosure	6
4.6 Finance	6
5 Key aspects of the Integrated Population Statistics System	6
5.1 An address register	6
5.2 A population register	6
5.3 A linked statistical database	6
5.4 A 2011 Census and Coverage Survey	6
5.5 An Integrated Population Statistics database	7
5.6 Beyond 2013	7
6 Next steps	8
Annex A: Key drivers behind the proposal	9
A.1 A traditional census	9
A.2 User requirements	9
A.3 Changes in society	9
A.4 Administrative data initiatives	10
A.5 A traditional census versus population estimates	10
Annex B: Links to other documents	11

Proposals for an Integrated Population Statistics System

1 Background

Within the field of population statistics, there is an increasing demand for more frequent and detailed information for small areas and population subgroups. At the same time, falling response rates to traditional surveys and censuses are presenting producers of population statistics with increasing challenges. New and innovative methods for providing such statistics are therefore necessary. The American response to these issues has been to move to a continuous population survey (the 'American Community Survey'), whilst the French have implemented a rolling census (see **Annex B** for more information). The vision for an integrated population statistics system described here is the start of the Office for National Statistics' (ONS) response to these issues in England and Wales.

The future population statistics system could bring together data from a number of sources to produce a population statistics database that contains person-level information covering the whole population of England and Wales. Such information would provide a picture of the population far superior to anything currently possible. This population statistics system could begin to deliver population information early in the next decade. For the interim, in the review *A Demographic Statistics Service for the 21st Century*, ONS has outlined an action plan that will lead to continuing improvements in the reliability of population statistics. More information on these developments can be found via the links in **Annex B**, but they are beyond the scope of this discussion paper.

It is important to note that the population statistics system described in this paper remains a vision at present. There are a number of significant risks that would need to be addressed before it could become reality, many of which are beyond the control of the ONS. These risks are discussed in **Section 4** of this paper. However, the potential benefits of the system, addressed in **Section 3**, are considerable, and show it to be a vision worth striving for.

Comments on this proposal are invited by 28th November 2003.

2 An Integrated Population Statistics System

The proposed integrated population statistics system combines census, survey, and administrative data¹, linked at individual person level, to create a single, comprehensive population statistics database, which is updated over time. This database would underpin all ONS population and social statistics, resulting in significantly improved, more consistent statistics for the government community, the Health Service, academia and the private sector.

This form of population statistics system would provide highly accurate small area statistics much more frequently and quickly than is possible at present. It would give a far better understanding of the groups who make up our society, and of issues such as health, crime and education and the relationships between them. These benefits would be improved planning, delivery and monitoring of services and better targeted resources.

The key elements of this proposal are that:

- a high quality address register, covering all properties in England and Wales is needed;
- a full census operation would be developed and tested, for implementation in 2011;
- in parallel, a linked statistical database would be created, combining administrative and survey data by linking at the individual person and household level;
- following the 2011 Census, the linked statistical and census databases would be combined to create a linked population statistics database;
- from 2013 onwards, the linked population statistics database would be updated using administrative records, survey data, the address register, and any future population register, with consideration given to how such data answer user requirements; and

¹ For the purposes of this report, administrative data are taken to be those primarily collected for administrative purposes. Many may also already be used for some statistical purposes.

- the linked population statistics database would form the basis of all future population statistics produced by the ONS, including populating the Neighbourhood Statistics system and underpinning the mid-year population estimates series.

More detail on the key aspects of the proposed system is given in **Section 5**. The key drivers behind the proposal are given in **Annex A**. Detailed information on the research underpinning the proposal can be found in the documents referenced in **Annex B**.

3 Benefits of the population statistics system

Through individual-level linkage, the proposed integrated population statistics system would be extremely powerful and flexible. Some key benefits are provided in this section.

3.1 Statistical benefits

3.1.1 More frequent and timely small area population statistics

At present, nationally consistent, small area population statistics are only available from the census. Between censuses users are faced with a choice between using out-of-date census information or locally maintained or administrative data which do not have the standards of census data in terms of quality, comparability or consistency. The population statistics system would allow small area statistics to be provided alongside the annual mid-year population estimates.

Through the improved use of administrative data, the production of UK population statistics would move towards the model used in a number of other European countries such as Denmark, where quarterly population statistics are available within eight weeks of the end of the quarter. It should eventually be possible for ONS to produce population estimates within a similar time scale under this proposed system.

3.1.2 More consistent population statistics

The population statistics system would form the basis of all future population statistics produced by the ONS. By sourcing all population statistics from this system, more consistent statistics would result.

The current system of a decennial census with population estimates in the intervening years can lead to sharp revisions of the population estimates in census years. Such step-changes should be significantly reduced under the proposed system.

The population statistics database would provide an individual-level sampling frame for social surveys, with associated improvements in sample representativeness. Linking survey data to the database would allow an accurate assessment of non-response characteristics. Through these advances the quality of survey-based estimates would improve and a greater degree of consistency between survey-based social statistics and population estimates would be possible.

3.1.3 More detailed population statistics

Through individual-level linkage the population statistics system would provide multivariate analysis across data from different sources. For example, health characteristics could be compared with academic qualifications in ways not possible currently.

The longitudinal nature of the database would provide detailed information on population flows, although the accuracy of this information would be dependent on how up-to-date the source data are. The system would not be able to address all issues associated with measuring international migration without substantial changes to immigration and emigration procedures, but a better understanding of both internal and external migration would be possible.

Individual-level linkage would provide an insight into coverage and quality issues for the statistical sources comprising the population statistics system. This would result in improvements in quality of each source, with associated benefits for all uses to which these data are put.

3.2 Benefits to society

3.2.1 Improved policy formulation and monitoring

The population statistics database would form a powerful support for evidence-based policy formulation and review. Frequent and comparable small-area statistics would allow policies to be targeted on small areas effectively and to be monitored accurately. The increased availability of multivariate statistics would allow a better understanding of a breadth of issues facing society, including education, crime, health and the relationships between them.

3.2.2 Improved service delivery

The availability of more frequent, timely and accurate population statistics for small areas and sub-populations would allow more effective targeting and reallocation of scarce resources, with associated benefits for the development, delivery and monitoring of public and commercial services.

3.2.3 Improved efficiency

Over time further statistical data would be merged with the population statistics database so that a wide range of statistical users could benefit from the high quality core data held within the population statistics system, and it would need to be collected only once.

There are a number of developments across government (see **Annex A**) that aim to make improved use of administrative and statistical data through data linkage and identification numbers. These initiatives are largely being carried out in isolation from each other. The potential benefits of a more co-ordinated strategy are substantial.

3.2.4 Reduced public burden

Through improved use of existing data the requirement to collect information directly from the public for statistical purposes would diminish. Thus the public burden of participating in surveys and complying with the census would be reduced.

4 Risks to the proposed Integrated Population Statistics System

There are a number of significant risks to the successful implementation of the integrated population statistics system. The key risks are discussed in this section.

4.1 The legality of linkage

Without reform of the legislation governing the sharing of data across government, the integrated population statistics system will not be possible. This vision will necessitate statutory safeguards to ensure confidentiality of individual records linked for statistical purposes far beyond what we currently have.

4.2 Political and public acceptability

Public debate is required on the issues surrounding the wide-spread linkage of personal data. Record linkage can occur at a number of levels:

- aggregate data linkage, for example via geography;
- statistical linkage, which matches records with similar characteristics but not necessarily the same individual or household; and
- exact linkage, which matches specific individuals and households.

Where the linkage matches information for named individuals, the greater the necessity for legal and political acceptance and the stronger

the obligation to protect from disclosure.

Aggregate linkage is currently possible and underpins the ONS Neighbourhood Statistics Service (see **Annex B**). However, to cross-tabulate information from different sources aggregate linkage may introduce bias and inconsistencies and incorrect inferences will result. For the full potential of the population statistics system to be realised, exact linkage would be required.

Linkage of specific individuals is starting to be attempted across government and the wider statistical community, but with little co-ordination at present. The ONS and wider statistical community should take this opportunity to lead the public debate, demonstrate the significant benefits that can be achieved by responsible data sharing and linkage, and coordinate initiatives currently underway. We need to understand what protection is needed by respondents and how trust varies with particular circumstances.

4.3 Quality and availability of administrative data

The integrated population statistics system is dependent on the quality and availability of administrative data. Administrative data are collected to varying standards and definitions. They are also collected primarily for purposes other than the statistical ones for which we wish to exploit them. Therefore, the definitions that determine whether individuals appear on the data sources, the classifications used and even the existence of the data source itself could change without reference to the impact this would have on the population statistics system.

The implementation of the proposed system would provide an opportunity to develop a framework around which definitions could be harmonised. However, the quality, coverage and availability of the data sources would still pose a risk that would need to be managed carefully.

4.4 Dependencies

The population statistics system as described here is based on five key areas of development (see **Section 5**). Should any of these projects fail to come to fruition the design of the system would need to be re-assessed, but this would not necessarily imply that the lesser vision that is achievable is not worth pursuing. For example, should an address register not emerge from the current initiatives in the public sector, a register could be created in the short term as part of the 2011 Census operation. If the population register does not proceed then statistical matching could be used.

4.5 Confidentiality and disclosure

The use of detailed multivariate data at small areas has an associated risk that information about an identifiable individual will be disclosed. Safeguarding the confidentiality of an individual's information has always been a legal obligation of the ONS. All necessary actions would be taken to protect the confidentiality of all data held in the integrated population statistics system.

There are two aspects to this risk; the actual risk of the disclosure of personal information and the perceived risk. Both are closely related to the public acceptance of data linkage and the integrated population statistics system. This risk is manageable within ONS, which will work to manage both aspects.

4.6 Finance

The population statistics system is dependent on adequate funds being made available for the implementation of each of the strands of work. ONS will produce a detailed business case for the system, which will include clear deliverables from the various stages towards delivery of the full system.

5 Key aspects of the Integrated Population Statistics System

In its fullest extent the integrated population statistics system pulls together five key areas of work already underway across government and within ONS:

- the creation of a high quality address register;
- the Registrar General's feasibility study into the development of a population register;
- administrative data linkage;
- the 2011 Census; and
- forming a continuous population survey from the existing range of public household surveys.

The key stages to delivering this system are discussed below.

5.1 An address register

A high quality address register, which contains information on all properties including communal establishments and non-residential properties, is needed. In addition to basic information on the characteristics of each property, such a register would hold key characteristics of the population associated with

it to support statistical needs, including turnover rate and multi-occupancy. The register would also link to sources of information that identify prospective addresses. Continual maintenance would be undertaken to ensure that the register remained a high quality, relevant source of address information at least for its statistical purposes. This register would be based on work led by the Office of the Deputy Prime Minister (ODPM).

5.2 A population register

A UK population register is currently proposed, which aims to create and maintain a high quality database of core personal details on all people lawfully resident within UK borders, UK nationals living overseas and deceased citizens where these latter records are necessary for the conduct of government business. The population register would store core personal details of name, address, date and place of birth, date of death and a unique personal identifier, and would provide a valuable spine for linkage of data from different sources. With explicit and special legal authority, such data may be available for statistical purposes, subject to public acceptability and the overseeing of its custody.

5.3 A linked statistical database

A linked statistical database would be created that links data from a number of sources, including administrative and survey data. By linking at an individual level, this database would enable cross tabulation of variables from different sources for small areas and populations. There would be significant legal hurdles to this aspect of the system to overcome, and public acceptability would need to be addressed. These issues are covered in more detail in **Section 4**. However, even if the legal and public acceptability issues were to be resolved in the near future, alternative data sources and systems are unlikely to be sufficiently developed to replace a census in England and Wales by 2011.

5.4 A 2011 Census and Coverage Survey

In parallel with the implementation of the linked source database, ONS are developing a census model for implementation in 2011. This model builds on lessons learned from the 2001 Census and on recent technological developments with an emphasis on maximising response rates and confidence in the final results (see **Annex B**).

A coverage survey will be used to assess the census results for under- and over-coverage. Additionally, data from the coverage survey would be matched to the linked statistical

database to give an alternative population estimate, which would be used to evaluate the census results.

5.5 An Integrated Population Statistics database

Following the 2011 Census, an integrated population statistics database would be created which links, at an individual level, the data collected by the 2011 Census, the linked statistical database and the coverage survey.

From 2013 onwards, the population statistics database would be updated regularly using information from administrative sources, the address register, the population register and surveys, creating a longitudinal database covering the whole population.

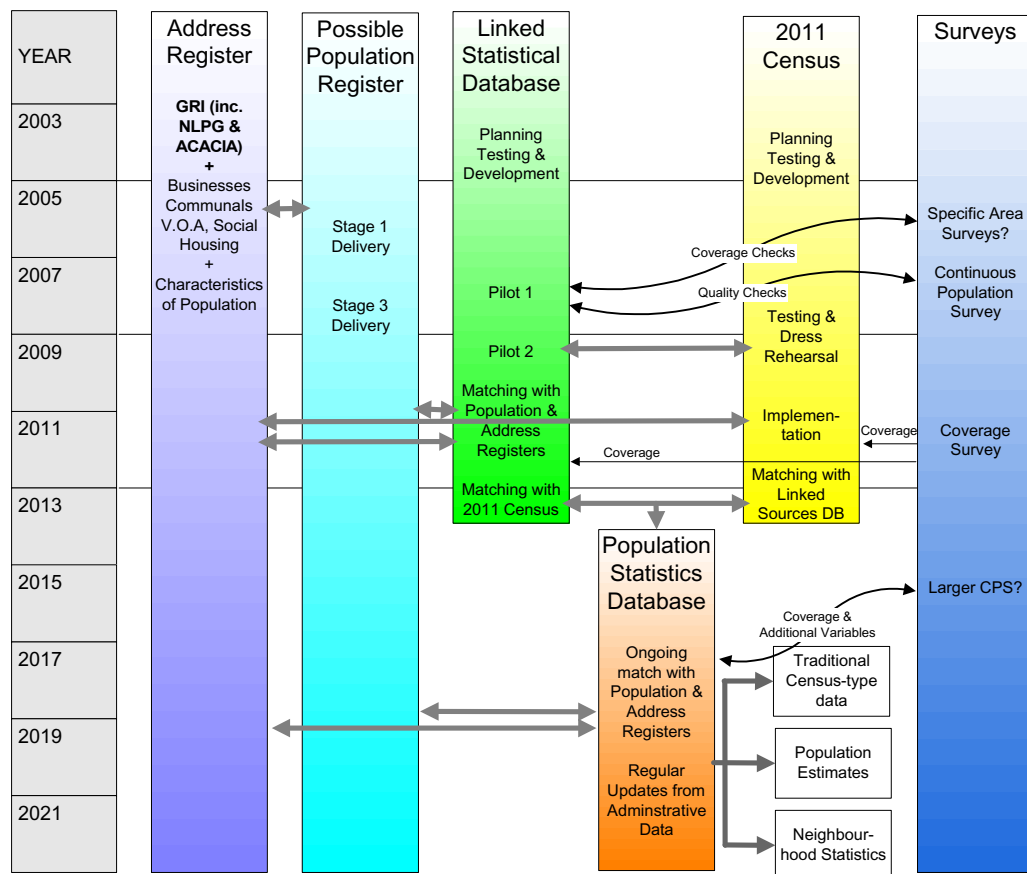
5.6 Beyond 2013

Consideration must be given to how to provide information required for key uses which is not available from administrative data. For 2011 this information would be collected as part of the Census. Beyond this point there are a number

of ways in which such information could be provided, including a full census in 2021, a rolling census, via the Continuous Population Survey or by expanding the information collected in administrative or registration systems. The appropriate mechanism for providing this information would be dependent on the frequency, accuracy and geography at which the information is then required and the extent of development of other data sources, most critically the population register.

Figure 1 illustrates the process by which the integrated population statistics system would be implemented.

Figure 1
Illustration of the proposed integrated population statistics system



6 Next steps

In order to advance this set of proposals, comments are invited. In particular, comments would be welcomed on:

- **Desirability.** From discussions with users to date there appears to be strong support for this proposal. **We would welcome confirmation of this view, or otherwise.**
- **Specific benefits.** Whilst we are confident that we have identified sufficient benefits to justify the cost of the integrated population statistics system (see **Section 3**), there may be significant benefits that we have not identified and which we would wish to include in our business case. **Users are welcome to indicate other such benefits.**
- **Risks.** Section 4 identifies a number of significant issues to be overcome for this vision to become a reality. **Users are invited to identify any others.**
- **Scope.** **We would welcome views on how the implementation of the population statistics system could be co-ordinated across government, the Health Service, academia and the private sector.**

Comments that are available by **28 November 2003** will be of help to ONS in shaping further developments. Please send them to:

Peter Benton
Census Strategic Development Programme
Office for National Statistics
Segensworth Road
Titchfield
Hampshire
PO15 5RR

email: csdp@ons.gov.uk

Annex A: Key drivers behind the proposal

This annex provides information on the considerations that gave rise to the proposal for an integrated population statistics system.

A.1 A traditional census

The census currently provides statistics relating to:

- population units: a benchmark count of people and housing, with key characteristics such as age and sex;
- population structures: information on households and families; and
- population characteristics: consistent and comparable data, with the range of topics allowing multivariate analysis, giving rich information down to small areas and population sub-groups.

There is no single alternative to a full census operation in 2011 that can deliver all three aspects of census data. However, traditional censuses across the world are facing the challenge of falling response rates, with specific subgroups becoming increasingly difficult to reach. Providing intercensal estimates of the required reliability for a variety of population bases is another challenge. Coupled with the difficulties of estimating undercount via surveys, this provides a strong incentive to investigate how other administrative and statistical data could improve on the quality of census data.

A.2 User requirements

It is clear from consultation with users that the requirement for the kind of information traditionally provided by a census will not only continue, but become greater and more complex. As statistical information is used to target policy delivery, as well as policy design and evaluation, its reliability needs to be known.

In a recent speech ‘The balance of power - enabling delivery and reform of public services’, Sir Andrew Turnbull, Secretary of the Cabinet and Head of the Home Civil Service, outlined the Prime Minister’s four key principles behind the reform of public services:

- high national standards with clear accountability;
- devolved decision-making to the frontline;
- greater flexibility in how, when and what services are delivered, including greater staff flexibility and the use of incentives; and
- more consumer choice - where appropriate, alternative providers should be offered.

Implicit in these principles is a requirement for more frequent, detailed, comparable information nationally, regionally and locally, for small areas and small populations.

A.3 Changes in society

Changes in society change user requirements for information. The increasing mobility of the population, both within the UK and internationally, mean that the state of the population is changing much more rapidly than in the past and we want to know more, more often and with more detail. Local populations vary by day of the week, seasonally, and through time, with complex geographical patterns, as people move for work, family and social reasons. Population structures are also increasingly fluid and complex. Individuals can reside in more than one address, household, family, geographic area and even country.

If policies and decisions are to be founded on reliable information then:

- a greater variety of population definitions are required, for example the concept of ‘usual residence’ needs to be clarified and supplemented by other definitions to account for the increasing occurrence of ‘second homes’, ‘weekday residences’ and residents who are temporarily within the local area or the UK; and
- more sophisticated and robust ways of assessing and assuring the quality of population counts are required.

A.4 Administrative data initiatives

A number of developments are underway which aim to make improved use of administrative data, including:

- Neighbourhood Statistics, which provides information for small areas across the UK;
- the development of a UK population register;
- the recent announcement of unique reference numbers for all children;
- the possible introduction of identity/entitlement cards; and
- many record linkage exercises, across government and the private sector.

These exercises are largely being carried out in isolation from each other; the potential benefits of a more co-ordinated strategic approach are substantial.

A.5 A traditional census versus population estimates

The need for more frequent information, and the potential to move away from the decennial census cycle towards annual information from administrative and statistical data removes the distinction between the census and other inter-censal population statistics. A combined approach is required so that the resulting system can provide measures of comparable, known reliability throughout the decade.

Annex B: Links to other documents

User Requirements Report
(www.statistics.gov.uk/statbase/Product.asp?vlnk=10784)

Alternatives to a census: review of international approaches
(www.statistics.gov.uk/statbase/Product.asp?vlnk=10784)

Alternatives to a census: review of previous UK studies
(www.statistics.gov.uk/statbase/Product.asp?vlnk=10784)

Alternatives to a census: Rolling Census report
(www.statistics.gov.uk/statbase/Product.asp?vlnk=10784)

Alternatives to a census: Administrative sources
(www.statistics.gov.uk/statbase/Product.asp?vlnk=10784)

The 2011 Census: Proposed Design
(www.statistics.gov.uk/statbase/Product.asp?vlnk=10790)

For more information about:

A demographics statistics service for the 21st Century
(www.statistics.gov.uk/about/Methodology_by_theme/Dem_Stat_Ser_21ST_Cen.asp)

Neighbourhood Statistics Service (NeSS)
(www.neighbourhood.statistics.gov.uk)

Citizens Information Project (CIP)
(www.statistics.gov.uk/registration/cip.asp)

National Land and Property Gazetteer
(www.nlpg.org.uk)

ACACIA project
(www.idea.gov.uk/press/acacia_report.pdf)

UK Continuous Population Survey (CPS), formerly known as the Integrated Social Survey (ISS)
(www.statistics.gov.uk/ssd/downloads/Paper_ISS_V2.pdf)

US American Community Survey (ACS)
(www.census.gov/acs/www/index.html)

French Rolling Census (in French)
(www.insee.fr/en/recensement/page_accueil_rp.htm)