

The Census and Future Provision of Population Statistics in England and Wales: Public attitudes to the use of personal data for official statistics

March 2014

For more information

Search Beyond 2011 @ www.ons.gov.uk or contact: beyond2011@ons.gov.uk

© Crown Copyright 2014

Public attitudes to the use of personal data for official statistics

The Census and Future Provision of Population Statistics in England and Wales: Public attitudes to the use of personal data for official statistics

Table of Contents

Glossary	2
1 Overview	3
2 Introduction	5
2.1 Background.....	5
2.2 Scope and content	6
3 Summary findings from each strand of research	7
3.1 Cognitive testing.....	7
3.2 Surveys	8
3.2.1 The 2009 survey	8
3.2.2 The 2012 survey	8
3.2.3 The 2013 survey	9
3.3 Focus groups and structured workshops	11
3.3.1 ONS focus groups.....	11
3.3.2 Public dialogue workshops	11
4 Conclusions	13
5 References	14

Glossary

Terms	Description
Administrative data	Data already collected by public bodies to provide a service or to meet an operational need, which can be re-used for statistical purposes.
Anonymised data	Data relating to a specific individual, where personal information has been removed to prevent identification.
Data	Qualitative or quantitative statements or numbers that are assumed to be factual, and not the product of analysis or interpretation.
Data linking	The creation of associations between data held in two or more different sources.
Data sharing	The transfer of administrative data between different organisations.
Dataset	A collection of information held in structured electronic form.
Disclosive	Data are potentially disclosive if, despite the removal of obvious personal identifiers, the characteristics of any dataset (in isolation or in conjunction with other datasets in the public domain) might lead to the identification of an individual.
Information	Output of some process that summarises, interprets or otherwise represents data to convey meaning.
Personal data (as defined by the Data Protection Act 1998)	Data relating to a specific individual where the individual is identified or identifiable in the hands of a recipient of the personal data.
Sensitive personal data (as defined by the Data Protection Act 1998)	Sensitive personal data means personal data consisting of information as to (a) the racial or ethnic origin of the data subject (b) his political opinions (c) his religious beliefs or other beliefs of a similar nature (d) whether he is a member of a trade union (within the meaning of the Trade Union and Labour Relations (Consolidation) Act 1992) (e) his physical or mental health or condition (f) his sexual life (g) the commission or alleged commission by him of any offence, or (h) any proceedings for any offence committed or alleged to have been committed by him, the disposal of such proceedings or the sentence of any court in such proceedings.
Statistical purposes	Purposes which involve the collection and processing of data necessary for producing statistics.

1 Overview

As part of the Beyond 2011 Programme of work on the future provision of population statistics and the census, the Office for National Statistics (ONS) has carried out research into public attitudes to the collection and use of data for production of official statistics and research, including different potential approaches for producing population statistics. This research was undertaken over the period 2009-2013 and included original quantitative and qualitative research undertaken by ONS, as well as research commissioned by ONS and undertaken by Ipsos MORI.

ONS also undertook an extensive review of literature from a range of sources including reports produced by government departments, national statistics institutes, academic organisations, think tanks, commercial research organisations and the media.

In summary, the research has demonstrated that:

- there is generally a very low level of public understanding about data, how it is collected and used, and only basic knowledge of the related vocabulary;
- the public generally does not understand the difference between operational and statistical uses of personal data;
- nearly half of the public assume that government already routinely links data about the population from multiple sources in a central data store;
- around three quarters of people do not object to data held by other government departments being shared with ONS;
- the public are supportive of data sharing when personal or public benefit can be demonstrated and these are communicated effectively;
- data linking and storage is more acceptable if the personal data are anonymised. However, explaining the process of anonymisation is complex and difficult for the public to understand;
- any objections to the use of personal data are largely related to security and privacy concerns;
- the public get most of their messages about the use of personal data from the media and their own personal experiences;
- around three quarters of the public trust ONS to protect the confidentiality of their data;
- the public is generally positive towards the decennial census as a means of gathering information about the population; and
- when provided with reassurance with regard to security and privacy, the public broadly support ONS re-using administrative data to produce statistics.

The public's views of the acceptability of the use of personal data for statistical and research purposes differ according to who is using the data and for what purpose. While three quarters of the public do not object to data being shared with ONS, around three quarters raised concerns about data sharing in general. These concerns encompass:

- security and confidentiality;
- privacy and anonymity;
- transparency, control, consent and trust;
- governance and regulation; and
- public and personal benefits.

Public acceptability of the use of data is improved by:

- appropriate communications, in particular, knowing to what use the data will be put and what benefits will result; and
- ensuring that safeguards relating to security and privacy are in place and communicated effectively.

The research has shown that public views are not static and will continue to evolve in the light of public information and debate. Further research is therefore required, both to understand and refine the results of the research presented here, and to monitor public views as they change over time.

It is important to emphasise that the issues discussed here are complicated and the findings are open to interpretation. ONS is presenting its own conclusions from ONS and other research, and will publish the detailed reports from each research exercise so that others can scrutinise the work and draw their own conclusions.

ONS welcomes further discussion of any of the issues raised.

2 Introduction

2.1 Background

Every ten years, for over 200 years, each household in England and Wales has been required to respond to the census. The statistics it provides underpin the allocation of billions of pounds of public spending each year and are fundamental to decision making, policy formulation and outcome monitoring across government, the commercial sector and the wider community.

The ONS Beyond 2011 Programme has been reviewing requirements for population statistics and how they might best be met in future. Improvements in technology and in government data sources offer opportunities to either modernise the existing census process, or to develop an alternative census method that re-uses existing data already held within government.

A number of approaches have been considered, and each has been assessed against a published set of evaluation criteria including statistical quality, cost, technical and legal feasibility, public acceptability and public burden. This work has resulted in two potential approaches for census taking in future on which ONS consulted in autumn 2013, as follows:

- once a decade, like that conducted in 2011, but primarily online;
- using existing government data and compulsory annual surveys.

Public acceptability has been identified as an important issue when considering making increased use of individuals' personal administrative data within the population statistics system (Jones and Elias, 2006; Dugmore et al, 2011). Since 2009 ONS has undertaken research designed to develop a comprehensive understanding of public attitudes relating to the re-use of administrative data for research and statistical purposes. ONS work has focused on:

- understanding of the terms and concepts related to data and statistics;
- understanding of how government holds and shares basic information;
- attitudes towards the use of personal data;
- attitudes about data linkage and the creation of linked datasets;
- attitudes towards the decennial census; and
- attitudes towards the production of statistics using existing administrative data sources.

It is important to note the distinction between the use of data for statistical and research purposes, and its use for operational purposes. In the case of statistical use, no information about an identifiable individual will be disclosed (the data are published in aggregated form) and no actions or decisions will be made which could have a direct individual impact. Operational uses generally require identifiable data about an individual for the provision of public services, and may impact directly on that individual – for example the use of information in administering the tax or benefit systems. This report focuses on the use of personal data for statistical or research purposes only.

What has become clear from the research is that the issues are complex, and the public's views change significantly depending on the context of the discussion. The findings of different parts of the research can therefore often appear contradictory. For example, the results of surveys can differ if different questions are asked in combination, or if the wording of questions changes between surveys. Views expressed in focus groups can be different again, as people's understanding develops through discussion of the issues. This report seeks to identify common themes and messages but the conclusions are open to interpretation. ONS welcomes further discussion on these issues.

2.2 Scope and content

The research conducted by, or on behalf of, ONS used three key methods:

- 'cognitive testing' of people's understanding of concepts and terms related to data and statistics (in depth one-to-one discussions with a small number of individuals);
- surveys covering nearly 5,000 people in total to assess public understanding of how the government holds and shares personal data, and attitudes towards data sharing for the purpose of statistics and research; and
- focus groups and structured workshops with small groups of people to explore issues and opinions in more depth.

The key findings of the separate research exercises are summarised in Section 3. Finally, Section 4 provides summary conclusions.

A list of all sources used is provided in Section 5. Some have been published already and the remainder will be published in April 2014. Section 5 indicates those which are already publicly available.

3 Summary findings from each strand of research

3.1 Cognitive testing

ONS commissioned Independent Social Research Ltd to undertake cognitive testing between January and May 2013. Interviews were held with 45 participants, in three waves, in five locations (Cardiff, London, Birmingham, Bristol and Elstree, Hertfordshire). The first two waves were designed to test public understanding of an agreed list of words, terms and concepts that might be used by ONS when discussing data sharing and the production of statistics. The third wave involved the testing of a set of survey questions on attitudes to the use of administrative data for statistical purposes, designed for use in the ONS Opinions and Lifestyle Survey.

The key finding from the research (ONS, 2014a) was that the public are not familiar with, and often have no understanding of, the terms and concepts commonly used in relation to the production of population statistics. There is generally a very low level of understanding about data, how it is collected and used, and only basic knowledge of the related vocabulary. Figure 1 below, shows that the public have a general understanding of commonly used terms such as ‘data’ and ‘statistics’ but have no understanding of more complex, technical terms such as ‘anonymisation’ and ‘disclosive’.

FIGURE 1: Public understanding of terms and concepts relating to the census approach and the administrative data and surveys approach

General understanding	Limited understanding	No understanding
<ul style="list-style-type: none"> data characteristics population record statistics 	<ul style="list-style-type: none"> address register dataset data sharing demographic mandatory socio-demographic 	<ul style="list-style-type: none"> administrative data anonymisation attributes coverage survey data linking data matching disclosive operational data unique identifier

Source: ONS, 2014a

3.2 Surveys

ONS has undertaken three sets of survey-based research since 2009, all using the ONS Opinions and Lifestyle Survey. Note that results of all three surveys have been adjusted to take account of bias that may have been caused by non-response.

3.2.1 The 2009 survey

The 2009 survey drew in part on an innovative 2004 survey, which was itself part of a bigger project considering the merits of a central database and data sharing for statistical and operational purposes (the Citizens Information Project). The 2009 survey had seven questions (four of which had already been asked in 2004). Re-using those questions in 2009 enabled an understanding of how public attitudes to data sharing had changed over time. The survey had a 53 per cent response rate with 974 respondents in total.

Key findings (ONS, 2009) were that:

- Over half of respondents were aware that no single government central population database currently existed, but that there were separate databases maintained by individual departments.
- Approximately two thirds of respondents said they were in favour of the Government creating a single central population database, for either statistical or operational purposes; one fifth opposed it and the remainder had no view.
- Those who were opposed to the creation of a single central population database raised concerns over data security, information leaking and too much information sharing across government.
- A majority of respondents believed that, if a central population database were created, name, address, date of birth, National Insurance number and Passport number should be included.
- Public views on the key issues remained very similar to the 2004 survey. In particular, support for the creation of a single central population database remained broadly at the same level (65 per cent in 2004 and 64 per cent in 2009).

3.2.2 The 2012 survey

The second survey, in February 2012, asked 12 questions covering the public's views and understanding of the census, on the sharing and holding of administrative data, and questions to enable a comparison with the findings of the 2009 survey. The survey had a 59 per cent response rate, with 1060 respondents in total.

Key findings (ONS, 2014b) were that:

- Overall, the reasons for census-taking are widely understood by the public, with 9 out of 10 people knowing that the census is used to understand the make-up of the population and to assess population change.

- Three quarters of people are happy to provide potentially sensitive personal information on a census form, although one in five object to giving information about their income¹ and one in ten do not want to provide information on their sexual orientation².
- People are more likely to agree with ONS obtaining personal data from other government departments than from commercial organisations (one half and one third respectively).
- Around three quarters of the public trust ONS to protect the confidentiality of their data.
- Just over half of the public support personal information being held in a database for statistical purposes and a further fifth neither agree nor disagree. Around one seventh of the public are opposed to the creation of a database for statistical purposes, raising concerns over personal privacy, data security and unauthorised access, as well as general objections to the amount of data being shared.
- If a statistical database was created, a large majority believe that personal details such as name, address, date of birth and sex should be included. Nearly half think that it would reduce costs and help to deliver services more effectively.
- Around two thirds of people thought that personal details such as ethnicity, marital status and employment status should be included.
- However, two thirds of people were opposed to information about income data being held on such a database.
- Comparison with the 2009 survey showed that support for a population database had reduced slightly (64 per cent in 2009 and 59 per cent in 2012).
- However, there was significantly stronger support for a wider range of personal information being held in 2012 than in 2009 – for example around two thirds of people thought that ethnicity, marital status and employment status should be included in a central population database, compared with around one third of people in 2009.

3.2.3 The 2013 survey

The third survey ran in three waves from May to July 2013, using the refined set of questions developed through the cognitive testing described above. It covered the public's views on the sharing and holding of administrative data, the anonymisation of data, and how views vary depending on the types of organisations involved. The survey had a 52 per cent response rate, with 2,737 respondents in total.

In addition, the sample was split into two parts. Two thirds of respondents were read a 'privacy pledge' reassuring them that ONS looks after the data it collects; the other third were not. The aim of this was to understand whether people's views were different if such reassurances were prominent in their minds.

¹ Note that the 2011 Census did not ask a question about income.

² Note that the 2011 Census did not ask a question about sexual orientation.

Key findings (ONS, 2014c) were that:

- Just over half of the public supported personal information collected by other government departments and public organisations being used by ONS for statistical purposes; around one third disagreed and around a fifth neither agreed nor disagreed.
- Support increased very slightly (by four percentage points) if the data were to be stored by ONS temporarily rather than permanently.
- If the data are anonymised (name, address and other personal identifiers removed) three fifths of the public supported personal information collected by other government departments and public organisations being used by ONS for statistical purposes. One quarter did not agree, and the remainder neither agreed nor disagreed.
- Respondents to whom the privacy pledge was read were six percentage points more likely to agree with ONS using personal information collected by other organisations, regardless of whether the data were anonymised.
- People's views about providing their personal details varied significantly depending on the type of organisation to which it was provided, and the nature of the information provided. The patterns were complex, but in general people were more willing to provide information to public sector organisations (central government, the NHS, local councils) than to private sector organisations (retailers or market research companies) with the exception of financial institutions.
- People were generally more supportive of ONS receiving personal data from public sector organisations than from private sector organisations, but again patterns were complex.

Questions in this survey were not consistent with those asked in the previous surveys so the findings cannot be compared precisely.

3.3 Focus groups and structured workshops

ONS conducted research via focus groups in 2010, and commissioned further workshop-based research by Ipsos MORI in 2013 (commissioned jointly with the Economic and Social Research Council (ESRC)). The structured workshops deployed a ‘public dialogue’ approach, meaning that information was presented in a structured way to inform the workshop participants, with experts on hand to discuss and describe technical issues as required.

3.3.1 ONS focus groups

Key findings (ONS, 2014d) were that:

- Participants did not have a clear understanding of what data were held by government, which government departments hold this information, what it is used for and who it is shared with. Some suggested that a central database may already exist, but most believed that there were a number of separate databases.
- It was clear that it was difficult for participants to translate data into statistics, and statistics into decisions, actions and ultimately benefits for the public. There was also a clear lack of understanding of the distinction between the use of data for statistical purposes and its use for operational purposes.
- Participants’ acceptance of data sharing had certain limitations, such as a clear reason why the data needed to be shared. There was an expectation that a person providing information to the Government would see a direct, beneficial effect from doing so.
- Participants felt that government departments did not communicate with the public regarding what data they hold, and what they use it for.
- Participants felt that the Government was incompetent in handling data, and this idea had been reinforced by negative stories in the media.
- Participants also felt that the Government did not take its responsibility for the public’s personal data seriously.
- There was some desire for the public to regain control of their data, and for the Government to ask permission to use it.
- Data security was a major concern for participants, most of whom would countenance a central population database only if appropriate security could be put in place.
- However most participants were reassured by current ONS processes for limiting access to data and safeguarding confidentiality.
- Reassurance about security is therefore key, and educating the public about how statistics are kept anonymous is important.

3.3.2 Public dialogue workshops

Between 6 October and 9 November 2013 Ipsos MORI conducted seven pairs of six-hour long public dialogue workshops in London, Manchester, King’s Lynn, Cardiff, Wrexham, Stirling and Belfast. Each workshop involved 16 to 20 participants (129 participants in total) and each participant attended a pair of workshops.

The purpose of the workshops was to consider a range of issues relating to the census, and also to the proposed development of an Administrative Data Research Network (ADRN) by the ESRC to support social research. During the workshops the process of anonymising data (described as ‘de-identifying’) was explained to the participants in detail, and their views about data sharing were discussed both with and without such de-identifying.

Key findings (Ipsos MORI, 2014a, 2014b) were that:

- Participants had little prior familiarity with, or understanding of, the purpose, processes and uses of social research or statistics.
- Participants found the operational and statistical/research uses of data difficult to separate conceptually.
- By the end of the second day of dialogue, most participants had concluded that they supported administrative data linking if three main conditions were met:
 - the data are fully de-identified;
 - the data are kept secure at all times; and
 - the data are linked for socially beneficial purposes.
- Confidence in the process of de-identification was crucial to creating support for linking administrative data. However, the findings from the dialogue demonstrated that the process was difficult and time-consuming to explain. In many cases, participants needed extensive information and discussions with experts and researchers in order to be satisfied that these conditions would be met.
- The census was not a key issue at the front of most people’s minds; consequently participants did not have strong feelings about the potential changes.
- The study suggested that there was openness to a change to the use of administrative data, where people had been given access to all of the arguments for and against change, had time to reflect, and time to ask questions about the potential changes.
- While there was an emotional attachment to the traditional census amongst some participants, overall the idea of making better use of data that has already been collected was seen as compelling.
- The possibility of cost-savings and more frequent (thus more up-to-date) data for decision-making resonated with participants.

4 Conclusions

Public attitudes research about the public acceptability of using personal data has shown that the public express mixed opinions about the use of their personal data for research and statistical purposes. The majority support the sharing of data with ONS for statistical purposes, but there are strong concerns about security and privacy which must be allayed.

The key issues in gaining public acceptance of the use of data are:

- appropriate communication, in particular, explaining what government is doing with data and the benefits that will result; and
- ensuring that safeguards relating to security and privacy are in place, and communicated effectively.

Given the general level of public understanding and the complexity of the issues, more research is needed to understand how the issues can be communicated in a manner which is understood by the majority of people. This needs to consider both how to express the key concepts, and the amount of information that is required. In particular, it will be important to communicate the difference between statistical and operational uses of personal data, and the security and privacy safeguards that are in place.

ONS will build on the results of this and previous work to enhance existing mechanisms of communication, and conduct further research and dialogue, before making detailed decisions about the future use of administrative data in the production of population statistics.

5 References

* denotes sources that will be published in April 2014

Dugmore, K., Furness, P., Leventhal, B. and Moy, C. (2011) *Beyond the 2011 Census in the United Kingdom* International Journal of Market Research 53 (5)

Ipsos MORI (2014a) *Dialogue on Data: Exploring the public's views on using linked administrative data for research purposes*

Ipsos MORI (2014b) *Dialogue on Data: Explaining the public's views on changes to the Census*

Jones, P. and Elias, P. (2006) *Administrative data as research resources: a selected audit*

Office for National Statistics (2009) *Data sharing between government departments: Report on public acceptability*

Office for National Statistics (2014a) *Beyond 2011 Programme Public attitudes research: Report of 2013 cognitive testing by Independent Social Research Ltd**

Office for National Statistics (2014b) *Beyond 2011 Programme Public attitudes research: Report on 2012 Opinions and Lifestyle Survey**

Office for National Statistics (2014c) *Beyond 2011 Programme Public attitudes research: Report on 2013 Opinions and Lifestyle Survey**

Office for National Statistics (2014d) *Beyond 2011 Programme Public attitudes research: Report on 2010 focus group research **