



Census 2001 Review and Evaluation

Origin Destination Evaluation Report

ONS is carrying out a review and evaluation of the 2001 Census in England and Wales which will culminate in a Data Quality report being published, followed by a General Report.

Plans for individual reports on specific aspects of the Census operation and a timetable for release have been published.

Each report is written in isolation and is subject to amendments as processing progresses and further information comes to light.

Reports will be released on the ONS website in the form of a high level Executive Summary and a more detailed Evaluation Report.

November 2005

Content	Page
Introduction	2
Consultation	2
Extraction of Data	3
Small Cell Adjustment.....	3
Release of Data	3
Use of Data	4
Conclusion / Possible Lessons for 2011	5

Census Customer Services
ONS
Titchfield
Fareham
Hants PO15 5RR

Telephone: ++44 (0) 1329 813800
Fax: ++44 (0) 1329 813587
Minicom: ++44 (0) 1329 813669
E-mail: census.customerservices@ons.gov.uk
Website: www.statistics.gov.uk/census2001

Census 2001 Review and Evaluation

Introduction

The reason for producing origin-destination (O-D) statistics was to provide customers with detailed data from the 2001 Census on migration and travel patterns. The O-D data differed from standard census output in one main respect. While, with standard output, the user could only specify one geographic area, with the O-D data there were two geographic areas to consider. The two areas in question were:

1. The area of origin. For migration data, this was the area containing the persons home address from one year before the Census while, for travel data, it was the area of the person's home address at Census time.
2. The destination area. For migration data, this was the area of the persons home address at Census time (i.e. where they had migrated to) while, for travel data, it was the area containing the address which they travelled to for work (or study in Scotland) at Census time.

The 2001 O-D data was produced, as a UK product, as a number of different tables at different geographic levels. The outlines of these tables can be found at : www.statistics.gov.uk/census2001/pdfs/migrants_final_layouts.pdf

The advantage of the O-D product was that it gave the user maximum flexibility in deciding what origin area and destination area they wanted. It also helped reduce the burden on the Census Offices as customers did not have the same need to request commissioned tables on migration and travel.

O-D data was something which had been produced for censuses prior to 2001. However, 2001 was the first time data on workplace/travel had covered 100% of census records. As a result of this, and the fact that Local Government reorganisation during the 1990s meant a lot of the 1991 Census data was for areas which no longer existed, the 2001 O-D data was keenly awaited. In addition, the use of the One Number Census (ONC) in the 2001 Census meant that, unlike 1991, the 2001 O-D data (in keeping with all other census data) covered **all** migration and workplace/travel flows.

Consultation

Consultation began in 1999 when the first firm proposals were put to users. As a result of comments received to these proposals, revised proposals were issued in 2000 for discussion at a series of workshops. Responses were received to these revised proposals and were summarised in two papers issued to the Census Advisory Groups. Further responses were received to these Advisory Group papers. The 3 Census Offices also received advice from user representatives on the Output, Geography and Confidentiality Working Group (OWG).

In September 2001, a booklet was published by the 3 Census offices entitled "Census 2001, Origin-Destination Statistics, Final Specifications", setting out the expected layout of each of the tables in the O-D product. Some further small changes were made to the layout of the tables after this booklet was published. In mid-2003, the stage had virtually been reached where all suggested changes from people outwith the 3 Census Offices had been considered and any necessary amendments made.

It had been hoped, at that point, to move forward with the production of the required tables, with a view to their possible release in late 2003 / early 2004. However, concerns were expressed, mainly by the Northern Ireland Statistics & Research Agency (NISRA) and the Office for National Statistics (ONS), about some aspects of the product. The main concerns were:

1. NISRA were concerned about Table W301 (workplace statistics by output area of residence and output area of workplace - the only output area level table for workplace statistics). Their main concern was whether the release of such data may breach confidentiality, by giving out information about the workforce of a particular employer, if there was one dominant employer in an output area. The duty of confidentiality in official statistics covers more than just the individuals filling in the return. ONS were sufficiently reassured though, with data quality issues such as respondent error and the amount of imputation which takes place in the Census, to release the data for this table for England & Wales. NISRA did, however, decide to exclude Northern Ireland data from table W301

Census 2001 Review and Evaluation

2. ONS raised a general concern as to whether there were any tables where the information in the tables ran the risk of releasing confidential information. ONS' Methodology Group did some work assessing this, with the eventual result that, in table W105 (workplace statistics by industry), some of the categories of industry were merged. The General Register Office for Scotland (GROS) decided to do likewise for table TV105 (travel Statistics for Scotland by industry).
3. NISRA were generally concerned at the effect which the Small Cell Adjustment Methodology (see later section) would have on some of the data and in particular (i) the impact that this could have on the utility of the data which was sparse in nature because of the large number of potential flows and (ii) the errors that would result if users were to aggregate the O-D data in order to obtain statistics for a higher level geography. This issue was discussed in detail by the Census Offices and the view was reached that any move to address these concerns would result in unacceptable further delays to the O-D product. It was subsequently decided that users would be alerted to the potential problems in the O-D supporting documentation and advised against making certain aggregations of the data. NISRA took the additional action of expanding the range of the Northern Ireland specific migration, travel to work and workplace population tables in order to negate the need for users in Northern Ireland to aggregate the O-D data.

Dealing with these issues meant that it was early 2004 before work started on producing the necessary tables. Because of the delay in starting work on this, and the effect this was going to have on release dates, it was decided to release the O-D product in five separate stages (as detailed in the later section "Release of data").

Extraction of data

The GROS took the lead responsibility for co-ordinating the O-D process through to release of the data. Once the final version of the table layouts was agreed, ONS extracted the raw census data, which was stored in a total of 112 different databases, one for each census Estimation Area. This clearly involved a significant amount of work for ONS.

Small Cell Adjustment

The Small Cell Adjustment Method was applied to all O-D data in England, Wales and Northern Ireland. It was also applied to the output area level Scottish travel data (table TV301).

After the data for the first O-D release (output area data) was produced, concern was expressed by ONS as to how the adjustment was working. It was first spotted that, when adjusted totals for the output area level data were produced at country level, and compared with unadjusted totals, the two sets of figures were more different from one another than had been expected.

Discussions then took place between the Census Offices and, although it was felt that the adjustment process was not working as randomly for the O-D data as the Census offices would have liked, it was decided that the remaining parts of the product should still be allowed to go out as planned. There were two main reasons for this:

1. The extra delay to the release of the O-D product, which would have inevitably resulted from reviewing and re-programming the adjustment methodology, was felt to be unacceptable, given that the product was already significantly later than had been originally planned.
2. It could not be ascertained to what extent rewriting the adjustment algorithm would actually improve the data.

Release of data

The O-D product was phased for release throughout 2004. It was released in the following order:

1. Output area level migration, travel and workplace data (May 2004).
2. Ward level migration (excluding moving groups), travel and workplace data (July 2004).
3. Postcode sector level migration (excluding moving groups) and travel data for Scotland only (July 2004).

Census 2001 Review and Evaluation

4. Local authority (Parliamentary constituency for Northern Ireland) level migration (excluding moving groups), travel and workplace data (October 2004).
5. Postcode sector (Scotland only), ward and local authority (Parliamentary constituency for Northern Ireland) level moving groups migration data (December 2004).

Use of data

CDs containing the raw O-D data were available to customers free of charge from any of the Census Offices. Although it was recommended that customers who wanted to analyse this data did so using a tailored software package, a large number of customers requested the raw data with a view to doing their own analyses. Most of the queries which came from such customers were fairly straightforward and easy to deal with. One or two customers, however, appeared to not fully appreciate that the raw data was not supplied in a format intended to be easily interrogated in a package such as Microsoft Excel.

There were two intermediaries who were able to provide versions of the data more readily analysed than products obtained directly from the Census Offices. These intermediaries were:

- The Greater London Authority (GLA), who manage the software package [SASPAC](#).
- The Census Interaction Data Service (CIDS) who operate the [WICID](#) (Web Interface to Census Interaction Data) system.

GLA and CIDS were asked for their views on both the positive and the negative aspects of the O-D product. The following were some of the views which were put forward by GLA and CIDS and their customers:

Positive aspects

- Users were satisfied with the design of tables and, in particular, the variables used to generate the (cross-tabulated) data.

- The quality of outputs was high compared with other census data outputs.
- The availability of data at output area level was welcomed, as it offered far greater spatial detail than had previously been available.
- The fact that all O-D data covered 100% of census records was seen as a major plus.
- The local authority level data was very well received.
- More tables and detail was available than in 1991.
- It was commented on that O-D was irreplaceable and very useful.

Negative aspects

- The later than expected delivery of the data.
- The effect of small cell adjustment, which had a greater impact on these tables than other census outputs, because of their sparse nature. Specific problems reported by users were:
 - Inconsistency in the data between different tables, when comparing both “double” and “single” geographies.
 - “Disappearing flows” i.e. flows between two areas appearing in some tables but not others.
 - The difficulty in quantifying the error involved in individual results.
 - Difficulty in producing estimated figures for non-standard geographies by aggregating up from output area.
- Less detailed aggregations (e.g. of age in table MG101) would have been preferred, due to the effect which small cell adjustment had on the data.
- The ward level data for England and Wales being for Census Area Statistics wards instead of the expected standard wards.

Census 2001 Review and Evaluation

- The lack of consistency in table design between different table geographies e.g. occupation data was available at ward level but not at local authority level.
- The age and ethnic groupings in the O-D data were inconsistent with a number of previously published “single-geography” tables.
- A more detailed breakdown of Standard Industrial Classification for travel to work data would have been welcomed.

Conclusion / possible lessons for 2011

1. The O-D product was successfully released before the end of 2004. The Census Offices worked very successfully together and the large distances between offices did not prove a barrier to getting the necessary work done. Despite this, from a customer point of view, it was regrettable that during the later half of 2003, there were delays in starting the work needed to produce the data. The knock-on effect this had on release dates is something which customers are hoping will not occur in 2011.
2. The impact of small cell adjustment on origin-destination statistics was greater than for other census outputs because of the sparse nature of the tables. This should be considered when discussing disclosure control methods for the 2011 Census.
3. The fact that ONS had to draw the required data out from 112 different databases was time consuming and affected the timescale within which the data could be publicly released. The Census Offices should consider whether a different database design could be used for the 2011 Census to alleviate this problem.

Census 2001 Review and Evaluation

Census Topics	Target Dates for Release
Legislation	Published
Non-Compliance (Executive Summary Only)	Published
Data Needs	Published
Geography	Published (Executive Summary)
Publicity	Published
Data Collection Development	Published
Data Collection Support	Published
Census Coverage Survey	Published
Processing	Published
Annex: Quality of Data Capture and Coding	Published
Downstream Processing	Published (Executive Summary)
Data Quality	
- Question non-response rates	Published
- Disclosure Control (Executive Summary only)	Published
- Data Validation (Executive Summary only)	Published
Edit & Imputation	Published
One Number Census	
- Quality Assurance	Published
- Lessons learnt (Executive Summary only)	Published
Output Policy	Published (Executive Summary)
Output Production	
- Part 1: Review of Output Released to date	Published (Executive Summary)
- Part 2: including Sample of Anonymised Records (SARs)/Origin Destination Matrices	Published
Census Access	Published
Programme Management	Published (Executive Summary)
Quality Report	Published
General Report	Published

Please note that the dates for release of individual evaluation reports noted above are target dates, and therefore subject to change. For the latest information please visit www.statistics.gov.uk/census2001/reviewevaluation.asp