



**Economic and Social
Council**

Distr.
General

ECE/CES/2006/33
3 April 2006

Original: ENGLISH

ECONOMIC COMMISSION FOR EUROPE

STATISTICAL COMMISSION

CONFERENCE OF EUROPEAN STATISTICIANS

Fifty-fourth plenary session

Paris, 13-15 June 2006

Item 6 of the provisional agenda

**SEMINAR ON POPULATION AND HOUSING CENSUSES
SESSION III**

Challenges and Changes in the England and Wales Census Design¹

Submitted by Office for National Statistics, United Kingdom

I. BACKGROUND

1. The design proposed for the 2011 Census for England and Wales results from a number of key factors:

- (a) experience from the 2001 Census, and preceding censuses, in England and Wales;
- (b) experience from other censuses around the world;
- (c) technological advances; and
- (d) changes in society.

2. In 2001, the England and Wales design included some significant changes from previous censuses. Firstly, respondents were required to return their completed questionnaires by posting them in pre-addressed envelopes, rather than enumerators collecting them after Census Day. The questionnaires were returned to 2,000 local offices from where enumerators collected and sorted them. Enumerators then carried out completion checks, following up those questionnaires that

¹ This paper has been prepared at the invitation of the secretariat.

were not fully complete and chasing those households from whom no questionnaire had been received.

3. Secondly, the capture of the data on the questionnaires was automated. Questionnaires were scanned and the information captured using various character recognition tools, supplemented by keying when required. This itself had a number of implications, the main one being that all of the information on all of the questionnaires was captured, rather than all information from just 10% of questionnaires, as in previous censuses.

4. Both of these design changes generally proved to be successful. However, the census was not without its problems. Post back was actually far more successful than anticipated, resulting in volumes of mail that exceeded the capacity of various parts of the postal operation. There were also problems with the field operation and with the automated data capture and coding operation.

5. A number of evaluations were carried out following the 2001 England and Wales census and these recognised both the strengths and weaknesses in the design and made recommendations for future operations. ONS has also assessed the methods implemented by other National Statistics Institutes (NSIs) in the 2000 and 2001 census round, as well as the proposed methodologies for those NSIs with censuses in 2006, to determine the benefits, or otherwise, that such approaches may offer to a 2011 England and Wales census. In addition, consideration was given to changes in technology and society and the influence that such changes may have on a future census.

6. The evaluations and recommendations, the findings from the assessments of other census operations and technology and societal changes, were pulled together and informed the proposed design.

7. The main factors underlying the proposed design are listed below.

(a) It is becoming increasingly difficult to recruit the numbers of enumerators required to carry out a traditional enumeration in some parts of England and Wales. A post back operation significantly reduced the numbers but there were still areas, particularly large metropolitan cities, where insufficient enumerators were recruited.

(b) Despite the fact that enumerators hand delivered questionnaires, in some areas thousands of addresses were missed. A number of factors combined to cause this problem. Firstly, the address list used was a number of years out of date. Secondly, the problems with recruiting resulted in many enumerators having significantly increased workloads and, thirdly, these increased workloads coincided with some of the areas that are the hardest to enumerate, in particular, the prevalence of multiple occupancy.

(c) The field management system failed. Headquarters (HQ) had little or no knowledge of how many questionnaires were delivered or returned nor how well, or badly, enumeration was going in any area. This meant that little or no action was taken for many of those areas where the enumeration was poorest, because information on progress was simply not available or became available when it was too late to act upon.

(d) The systems for the capture and coding of the questionnaires had not been fully tested before the start of the actual operation because procurement of the supplier did not start until after the 1997 Test.

(e) Generally, it is becoming harder to enumerate people. Peoples' lifestyles have changed; they are more mobile, less frequently at home and, for increasing numbers, it is much more difficult to define home. Also, people are becoming increasingly wary of information collected about them, more people are disillusioned and more are becoming disenfranchised and less willing to co-operate with Government initiatives. The concept of 'civic duty' is diminishing.

8. The main elements of the design proposed for 2011, some of which are to be tested in the 2007 Test, are:

- (a) census questionnaires will be posted out as well as delivered by enumerators;
- (b) respondents will have a choice of response channels; paper, Internet and, possibly, telephone;
- (c) questionnaires returned by post will be delivered to the data capture warehouse(s) through the postal channel and not to enumerators for checking;
- (d) checks for blank returns and partially complete returns will be carried out during the scanning operation;
- (e) the field force will carry out the follow up of non-response and blank returns.
- (f) Contact centre operators will carry out the follow up of partially complete responses by telephone;
- (g) Field staff will be targeted into the areas with the poorest response rates;
- (h) All data on all returns will be captured, coded and processed.

9. Some of the proposed elements of the design have been implemented in previous censuses in England and Wales and to an extent are tried and tested, although they may be impacted by changes elsewhere in the design. Others are new to England and Wales but have been implemented by census offices elsewhere, and England and Wales can benefit from these experiences. Others are, to the best of our knowledge, new to census operations and, down the line, our experiences may be of interest to other census offices. In addition, some of the elements, both new and tried and tested, require fundamentally different approaches to the supporting infrastructure and management support functions in order to enable the census to proceed. For example, offering internet capture, which is new to the England and Wales census, requires a robust reconciliation process with paper based responses and any other response channels.

10. The proposed design aims to achieve a number of, sometimes conflicting, objectives. The primary aim is to deliver a census that produces data that meets the users/ customers requirements. However, it aims to do this while minimising risk and maximising cost effectiveness. It therefore presents ONS with a number of challenges.

- (a) The England and Wales census counts households, individuals in communal establishments and individuals with no permanent address. To post out questionnaires, all the addresses and communal establishments requiring a census questionnaire need to be known prior to despatch for postal delivery. Individuals with no permanent address will be enumerated by hand.
- (b) A centralised post back strategy and targeting the field staff to areas of poor response, requires each questionnaire to be uniquely identified and an ability to track the whereabouts of each questionnaire.
- (c) The multiple response channels and reconciliation process required to support these.

II. CHALLENGES

A. An Address/ Household List

11. There have been a number of recent initiatives to develop an accurate, national address list between various Government bodies but, to date, these have been unsuccessful. However, Census not only requires a list of addresses, but a list of address with attributes, including identification of households and communal establishments, which is even less likely to be available. ONS is working with others, including Ordnance Survey and Local Authorities, to develop a national address register for the 2011 Census.

12. Traditionally, the addressing information used in the census has been verified or determined by the enumerator at the point of delivery. In 2001, each enumerator had a list of addresses for which they were responsible and had to manually check the status of each and update their record book with new information, as well as any addresses missed. Despite this process, in some places several thousand addresses and households were missed. A post out delivery requires that this information be available beforehand.

13. Work is currently underway to develop the address and household database that is required to support the proposed delivery strategy for the 2007 Test and two sources of addressing data are being evaluated. In September 2006, ONS will carry out an address checking exercise in each of the five areas of the 2007 Test. The Address Checkers will be issued with maps and a list of addresses to be checked within a defined boundary.

14. Two approaches are being implemented in the 2007 Test, which were developed following small scale tests carried out in 2005. In areas where the available information suggests little change in the address information, e.g., minimal new build, multiple occupancy or change of status, etc, the Address Checkers will visually check addresses. However, where they suspect multiple occupancy (because, for example, there is more than one door bell or satellite receiver), they will attempt to make contact to resolve any query. In other areas, mainly inner cities where high levels of multiple occupancy are anticipated, all addresses will be contacted to confirm details.

15. For both methods, there will be three attempts to make contact to obtain the information required. Address Checkers will also record additional address and any change to premises status. Team Managers will quality assure the work of Address Checkers in the field. For the actual census, ONS will also carry out an address checking exercise, which is currently scheduled for about September/ October 2010. Although work to develop a national address register is underway, ONS needs to plan for what it knows will be available for the census. At the moment, this does not include a national address register, including households, and, therefore, ONS has planned a national address checking exercise. Should an address register be available, of appropriate quality, then ONS may not carry out a national checking exercise, although there will be some checking undertaken. The amount of this will be determined by the factors underlying the requirement. If an appropriate quality address register is available nationally, the checking exercise may be random areas for QA plus specific areas identified as undergoing rapid

change or with high levels of multiple occupancy. An example of the former would be the east side of London because of preparations for the 2012 Olympics.

16. If the 2007 Test indicates that a post out delivery strategy does not detrimentally affect response rates and the operation runs smoothly, a post out strategy is likely to be adopted for the 2011 Census. The quality of the address register available at that time, will be one of the factors that informs where questionnaires will be posted to respondents and where hand delivery will be used.

B. Questionnaire Tracking

17. In previous censuses in England and Wales, enumerators have both delivered and collected completed questionnaires, from either the individual addresses assigned to them, or, as in 2001, from Royal Mail who collected the posted returns. In both cases, they checked the questionnaires for their level of completion and followed up addresses from which no, or insufficient, questionnaires (if hidden households had been identified) had been received or where the level of completion was unacceptable.

18. The post back design introduced in 2001 required respondents to post their completed questionnaire using the pre-addressed envelope. Royal Mail collected the questionnaires, sorted them to the 2,000 or so delivery addresses, and delivered them to the appropriate local area manager. Post back was far more successful than anticipated with about 88% of households returning their questionnaire through this channel; the estimate had been that about 70% would choose to do so. However, it was not without its problems.

19. The volumes of mail caused a number of problems for both the Royal Mail and the census operation. Because the volumes could not be managed, a back log of unsorted mail developed in a number of Royal Mail sorting offices, resulting in few deliveries of completed responses to the field staff. As a result, in many areas, enumerators were following up questionnaires that had already been returned. In some areas, field staff retrieved the returned questionnaires from the Royal Mail in order to carry out the sort themselves, so that non-response could be assessed and partial completion checks carried out and follow up instigated.

20. The 2011 design proposes a central postal return. This removes the need for a sort of the mail, which caused such a bottleneck in 2001. However, as a result, field staff will not know which questionnaires have not been received and therefore where to follow up.

21. In order to understand which questionnaires have not been returned, there needs to be a definitive list of all of the questionnaires delivered. Linking the delivered questionnaires to a specific address/household enables the list of non-responses to be drawn from the address list, once the responses received have been identified on the list. Therefore, every questionnaire needs to be uniquely identified and the unique identity number linked to an address.

22. In 2011, it is envisaged that most questionnaires will be pre-printed with an address and the link between the unique identity number and address established at this point. This process will be carried out by the printer during the print run. The printer then needs to pass the updated

database, which contains the link between the address and unique identity number back to ONS. However, there will be some unaddressed questionnaires and the link between the address and the unique identity number on these questionnaires needs to be established. It is proposed that this will be either by the back office when fulfilling a request for an additional, replacement or alternative questionnaire, or at the point of delivery by an enumerator. In either case, the link between the address and identity number needs to be recorded and the main database updated.

23. In order to update the database with responses as they are received, each questionnaire returned needs to be identified. In 2001, there were 8 days with over 1,000,000 questionnaires returned through the post, and on one of these, there were over 2,000,000. Each of the questionnaires needs to be receipted and the database updated to reflect the response to enable follow up for non-response.

24. Once non-response follow up has begun, probably a few days after Census day, the updated information on questionnaires received needs to be sent to the field in a way that minimises the follow up of questionnaires already returned. This means that the questionnaires requiring follow up need to be grouped and each group linked to a field manager. The information relevant to a specific manager then needs to get to that manager in a timeframe that maximises its usefulness.

The information gathered during the receipting process, together with the information on the questionnaires delivered, will enable an overall picture to be built up of areas where response is good and areas where response is poor. As the census progresses, the pattern of response rates will inform the distribution of work at the enumerator level and may inform HQ of major issues in specific areas.

25. To enable field managers to most effectively manage their teams, they will need access to and be able to analyse the information, or receive specific reports, on response rates within their area. This information may identify parts of their area with significantly lower response rates than other parts and therefore enable them to change the relative workloads of their enumerators to try and improve those areas with the poorest responses.

26. Census HQ need to understand the pattern of responses nationally and use the information to identify specific areas experiencing difficulty. Such information in 2001 may have enabled the introduction of strategies to overcome the under-enumeration in a number of cities. Should a significant problem be identified in an area, it will be possible to put in place a plan to address it and this may include moving enumerators from an adjacent area with a good level of response.

C. Multiple Response Channels

27. ONS is proposing to offer Internet capture in the 2011 Census. There are a number of challenges associated with this. The first is developing an Internet questionnaire that is robust, that reflects the paper questionnaire and is not difficult or unwieldy to complete. This includes ensuring that the questionnaire is accessible to all those who wish to respond via this channel, including the visually impaired, and that the information is secure.

28. The second is to ensure that there is an effective reconciliation process. This needs to be able to reconcile duplicate responses, i.e., a response made via the internet with a response made

on paper, and link an individual or continuation questionnaire completed on the internet with a household questionnaire completed on paper, and vice versa. ONS needs to know if one member of a household returned a completed questionnaire for the household via the Internet and another member of the same household completed a paper questionnaire for the household.

29. Although ONS is considering increasing the number of individuals on a household questionnaire from 5 to 6, approximately half a percent of households have more than 6 people. A household choosing to complete their questionnaire only using the internet will not require any additional questionnaire as the questionnaire on the internet will reflect the number of people identified as being part of that household. However, for those completing a paper questionnaire a continuation questionnaire will be needed. It is possible that the person completing the household questionnaire or the individual concerned, may choose to use the Internet. Where this happens, the reconciliation process must be able to establish the link between the continuation questionnaire and the household questionnaire to which it relates. The same applies to those members of a household who choose to complete an individual questionnaire, although in this situation it is also possible that an individual may choose to complete a paper based questionnaire when the household questionnaire was completed electronically. The same also applies to communal establishments.

30. The third challenge is the management of the volumes through each of the response channels and ensuring that each is appropriately scaled up to meet the overall expected demand as well as the demand profile. This presents a number of significant issues. It is very difficult now to estimate the volume likely to be received via each channel. Response to current Government initiatives using the Internet channel is about 11% but this will grow over the next few years. At this level, it is not appropriate to reduce the paper operation. However, there is a level of response beyond which ONS will need to scale back the paper operation because of the cost implications of running parallel activities.

31. Current work suggests that this point would be when over 40% of responses were expected through the Internet. To scale back the paper operation would require a pre-registration process. Up to half a percent of households move every month in England and Wales and careful track of whether these were registered for paper or Internet would need to be maintained to ensure that everyone requiring a paper questionnaire received one.

32. Also, scaling back the paper operation introduces significant risks, which would need to be managed. These include the time required to scale up should an event (e.g., the breach of a confidential site, such as a bank) significantly alter the expected volumes through the Internet channel. It takes about a week to print 2 – 3 million questionnaires, if the operation were scaled back by 30%, that represents 10,000,000 questionnaires and 3 to 4 weeks printing time. It will also take time to scale up the data capture and warehousing operation, assuming that the facilities already in place have the basic capacity! Either additional hardware would need to be procured and installed or the timetable for the operation increased. Additional staff would be required for the correction and coding and additional storage for the questionnaires themselves. In 2001, the questionnaires required 40 miles of shelving in an environmentally controlled warehouse. Increasing the capacity of an appropriate storage facility from 25 miles to 40 miles of shelving is not insignificant.

33. Similarly, the Internet capture facility also needs to be scaleable. It needs to deal with significant volumes in a very short time frame, without falling over or functioning so slowly that respondents give up.

III. SUMMARY

34. The design for England and Wales introduces a number of new ideas and methods. These have been introduced to overcome problems and risks identified in 2001 and previous censuses as well as to respond to the changing nature of society and how technology supports everyday life.

35. Embedded within the new design are also a number of challenges and risks. The three challenges identified were:

- (a) the need for a comprehensive address list that includes attributes to identify all households and communal establishments;
- (b) the need to be able to track questionnaires so the field force can follow up non response and ONS understand how the census is progressing across the country;
- (c) the introduction of alternative response channels and ensuring that all responses can be reconciled with related responses as well as the management of responses through each of the channels.

* * * * *