Chapter 6

Costs of retaining census records

Costs of retaining name-identified census records from one census were estimated by the Department of the Treasury in consultation with ABS and the Australian Archives.

The focus was on costs related to procedures for retention and storage of name-identified records. The potential costs of providing access to the retained records for research purposes in the future were not included. Operational decisions about what access to the records might be provided, and on what basis, would need to be made before the cost of this function could be assessed.

Three options for storing name-identified records were examined: storage of original paper forms, storage of microfilm copies of original paper forms and storage of information taken from the forms in electronic format.

The Committee does not consider the storage of the original paper forms to be economic. The Department of the Treasury estimated that the storage of microfilm copies of original census forms from the next census will cost $22.45 million. The Department of the Treasury estimated that storage of information from the forms in electronic format in the next census will cost $14.5 million. The Committee considers that the estimates provided are at the upper end of the cost spectrum.

Introduction

6.1 This chapter examines the cost of retention of name-identified records from one census.

6.2 Estimates were provided by the Department of the Treasury in consultation with the Australian Bureau of Statistics (ABS) and the Australian Archives. The focus was on costs related to procedures for

retention and storage of name-identified records which would be incurred by ABS and Australian Archives.

6.3 The Committee did not assess the potential costs of providing access to the retained records for research purposes in the future. Decisions about what access to the records might be provided, and on what basis, would need to be made before the cost of this function could be assessed.

6.4 In the absence of detailed specifications for some of the processes involved, cost estimates indicated only a general order of magnitude. All estimates were in current prices and were based on the estimated quantity of household and personal forms to be used in the 2001 Census. This is estimated to be around 8.8 million forms.[1]

6.5 Although the questions to be asked in the 2001 Census have not been finalised, the Treasury expects the number of questions to be similar to that of recent censuses. For costing purposes it was assumed that the number and sizes of pages in the forms would remain the same.[2]

Options for retention

6.6 Australian Archives stated that there are three options for storing name-identified census records:

- standard storage of the original paper forms by the Archives

1 Department of the Treasury, *Submissions*, p. S858.

2 Department of the Treasury, *Submissions*, p. S858.

- the making and storing by the Archives of microfilm copies, and, after appropriate checks, the authorised destruction of the original paper forms, and
- the retention of the records in electronic format (held by ABS or by the Archives) and the authorised destruction of the original paper forms.[3]

6.7 The option of retention of original paper forms is not considered to be feasible. Forms from one census would take up about 17 kilometres of shelving[4] and storage costs would be substantial.[5] More fundamentally, however, completed census forms returned under existing procedures could not be conserved reliably.[6] Substantial costs would be incurred in trying to preserve paper forms indefinitely.

6.8 The remaining two options for storage of the records were examined in some detail by the Committee.

The option of microfilms

6.9 The Australian Archives has stated that one of the most cost efficient and practical methods to preserve census records would be to film them, produce a high quality master and several duplicating copies, and destroy the original paper forms.[7]

---

3 Australian Archives, *Submissions*, p. S409.

4 Department of the Treasury, *Submissions*, p. S859.

5 Australian Archives, *Submissions*, p. S410.

6 Department of the Treasury, *Submissions*, p. S859.

7 Australian Archives, *Submissions*, p. S410.

6.10 Microfilm is the preferred archival preservation method internationally. The Treasury has commented that microfilming raises only simple reading and storage technology; it has a proven track record; and it does not pose serious problems of standardisation or obsolescence.[8]

6.11 The estimated cost of microfilming records from the next census is $22.45 million.[9] This estimate includes the costs of storage of the records during filming, preparation for filming and the microfilming itself, quality control checks after filming, storage at Australian Archives, address indexing, and costs incurred in preparation for and the conduct of the census to facilitate filming of the records.

6.12 The annual costs of storage (which are negligible) are not included in this sum. Nor does it include the costs of providing any future access to the retained records.

6.13 The following section reviews the components of the estimated cost.

8 Department of the Treasury, *Submissions*, p. S858.

9 Department of the Treasury, *Submissions*, p. S863.

A microfilmed copy of the entire census forms would easily fit into these cabinets

Storage during microfilming

6.14 Census forms are currently stored for up to 18 months while the statistical data are being extracted at the processing centre (located in Sydney for Census 1996).

6.15 The Treasury advised the Committee that if the forms were to be microfilmed before destruction, and the filming were conducted on-site, further storage and associated costs would be incurred. A commercial microfilming bureau approached to provide estimates of filming costs indicated that the filming process would take an additional 12 months.[10] The cost of an additional year of storage during filming is estimated to be $2.6 million.[11]

Project management and preparation for microfilming

6.16 Contract management, supervision, 'trouble shooting' and a range of physical preparation tasks are estimated to be $0.16 million.[12]

Microfilming

6.17 The Treasury stated that estimates for microfilming that had been sought from commercial microfilm bureaux came in at $1.83 million, $2.92 million and $9.74 million. The Treasury emphasised that these are only very rough indicative figures. One reason for the substantial variations between the estimates is differences in the

10 Department of the Treasury, *Submissions*, p. S859.

11 Department of the Treasury, *Submissions*, p. S860.

12 Department of the Treasury, *Submissions*, p. S860.

capabilities of their equipment but, the more important factor is differences in the standards of the final product they would provide.[13]

6.18 The Treasury considers that it would be important to set appropriate standards of quality in relation to such aspects as clarity, treatment of splices, and chemical stability and protection against film degradation. These standards would ensure the accuracy, legibility and durability of the copies. Because the original forms would be destroyed and the microfilm copies would be expected to last over one hundred years, quality standards would need to be much higher than for more standard microfilming. Most archives and libraries consider that the quality provided by the microfiche format is inadequate for preservation microfilming of handwriting.[14]

6.19 The Treasury also stated that a major consideration in setting quality control standards, of a different kind from the legibility and durability of microfilmed records, is the acceptable level of errors and omissions. The acceptable level would depend on the intended final use:

> researchers wishing to establish large-scale statistical patterns would not be greatly hampered if a very occasional form had not been properly microfilmed because unlike non-completion of census forms, faulty microfilming would result in a random loss of information and would not introduce systematic bias to the information remaining. On the other hand, a genealogist wishing to trace a particular person would clearly be frustrated if the census form for that person had been microfilmed incorrectly or not at all.[15]

13 Department of the Treasury, *Submissions*, p. S860.

14 Department of the Treasury, *Submissions*, p. S860.

15 Department of the Treasury, *Submissions*, pp. S860–S861.

6.20 The Treasury emphasised that the detailed specifications set for tenders would need to set very high standards to guard against loss of information and to retain the information as a resource without limiting the range of possible uses in 100 years time. On this basis, although the highest of the estimates, at $9.74 million, exceeds the other two, it better reflects the likely standards that would be required.[16]

6.21 The Committee agreed that high quality standards would be required to ensure that the records were microfilmed properly.

6.22 The Committee was concerned by the prospect that the project of microfilming census records could be outsourced to a non-government film bureau. The Committee holds the view that the use of a non-government bureau to film the records could raise doubts in the minds of the general public about the confidentiality of their census records. The Committee believes that no matter how ill-founded this perception might be, some public concern would be likely. The confidentiality of the records can best and most clearly be demonstrated to the public by ensuring the records are only seen and handled by staff of Australian Archives or ABS.

Checks after filming

6.23 The Treasury advised that standard preservation microfilming requires two kinds of post-production checking: a visual comparison of the film with the original to ensure that it has been copied correctly, and tests for chemical stability.[17]

16 Department of the Treasury, *Submissions*, p. S861.

17 Department of the Treasury, *Submissions*, p. S861.

6.24 The Treasury stated that in preservation microfilming a visual frame-by-frame check should be made against the original. This check is necessary because with automatic filming processes, misfeeds can cause skewing and stretching of film or can cut off images. There can also be missed documents, density and resolution problems, and damage to the film itself. The Australian Archives' practice for preservation filming was to require a frame-by-frame check before Archives is prepared to certify under the Archives Act that the film is a true and accurate copy.[18]

6.25 Tests for chemical stability would also be required. Checking one in every fifty rolls is common in large projects.[19]

6.26 The Treasury argued that the importance of the project would justify the use of a process called 'Silverlock' sulphide toning, which protects silver halide film from chemical degradation.[20]

6.27 In total, the costs of checks after filming are estimated to be $0.23 million.

Storage by Australian Archives

6.28 The costs of storing preservation standard microfilm in appropriate containers in an environment with strictly controlled

18 Department of the Treasury, *Submissions*, p. S861.

19 Department of the Treasury, *Submissions*, p. S861.

20 Department of the Treasury, *Submissions*, p. S861.

temperature and humidity are estimated to be small, at $7,000 annually.[21]

Indexing

6.29 Census forms are not currently indexed. During processing, they are held in packs according to census collection district.[22]

6.30 The Treasury stated that uses of the records that depend on looking up the copies of forms by census collection district would be facilitated by microfilming records in batches corresponding to collection districts and indexing them so that addresses of census respondents can be matched to collection district. The cost of address indexing is estimated to be $2.67 million.

Additional costs in preparation for and conduct of the census

6.31 The Treasury considers that additional costs would be incurred in training field staff and in extra time needed by staff to answer questions when leaving or picking up the forms. Training costs, at a rate of thirty minutes extra training per staff member, are estimated to be $0.3 million. Salary costs, at a rate of 5 minutes spent at every second household, are estimated to be $3.75 million.[23]

6.32 The Treasury argued that it would be incumbent on ABS to extend its public relations campaign to explain a change in policy relating

21 Department of the Treasury, *Submissions*, p. S862.

22 Department of the Treasury, *Submissions*, p. S862.

23 Department of the Treasury, *Submissions*, p. S863.

to retaining census forms and to keep public cooperation and response rates high. This is estimated to cost $3 million.[24]

6.33 The Committee considers that the costings provided in the Treasury submission are at the upper end of the scale of costs.

The option of electronic format

6.34 It is possible that information from the census forms could be retained in electronic format. Electronic format would provide researchers with a greater degree of flexibility in using the records, and most of the information from the census forms is already recorded in this format by ABS during processing.

6.35 The outcome of the census at the moment is a file of coded records for each person, family, household and dwelling enumerated in the census. The file contains no personal or household identifiers.[25] Names and addresses are not currently recorded but could be recorded along with the remainder of the information taken from each census form.

6.36 The Committee formed the view that under this option, the most suitable approach would be compilation of an electronic record of names and addresses by ABS. This record would be held separately from the information currently taken from the forms. The information could be matched in the future using a cross-referencing system which would

24 Department of the Treasury, *Submissions*, p. S863.

25 Australian Bureau of Statistics, *Submissions*, p. S354.

relate census information from one electronic file to the corresponding names and addresses on another electronic file.

6.37 The Treasury stated that recording names and addresses during statistical processing would delay the release of results from the census by three months.[26] The Committee believes that even if there were a delay of three months this could be tolerated. The Committee expects that such a delay would be shortlived, however, as improved processing techniques are expected to be introduced in the future due to technological advancement. Optical Character Recognition is one such new processing technology, which Dr Sui-Ming Tam of the ABS advised the Committee was under consideration for the 2001 Census. Furthermore, recording names and addresses during statistical processing would avoid delays due to double handling, which might otherwise occur.

6.38 One option is to use keyboard operators to key in the information from the forms. This option involves full manual data entry and verification of the data.

6.39 Optical Character Recognition (OCR) technology which scans the information, is a less costly option. OCR automates the initial capture of information from the forms, which makes this part of the process faster and less costly.

6.40 The Treasury emphasised that verification of the data is required to ensure that the information has been recorded correctly. The Treasury expects that the use of OCR to record hand-printed names and

26 Department of the Treasury, *Submissions*, p. S866.

addresses could lead to high error rates. To achieve a high level of reliability, key data entry verification by keyboard operators is required to confirm that the OCR process has correctly recorded the information from the forms and to rectify any errors.[27]

6.41 The estimated cost of electronic recording of names and addresses using OCR is $14.5 million.[28] This includes the costs of data entry for verification, repair of OCR responses, supervision of data entry, lease and support of work stations, accommodation, and programming and additional costs in preparation for and conduct of the census.

6.42 The Treasury also commented that electronic records need to be transferred periodically from their initial form to forms of storage which are technologically more advanced. This process is known as data conversion. It is necessary because of the rapidity with which hardware and software become obsolete – the records must be in a form which is compatible with equipment in general use.[29] The estimated cost of data conversion and of storage of the records is estimated to be an additional $20,000 per annum.

6.43 The Committee considers that the estimates provided by the Treasury are on the high end of the cost spectrum.

27 Department of the Treasury, *Submissions*, p. S866.

28 Department of the Treasury, *Submissions*, p. S866.

29 Department of the Treasury, *Submissions*, pp. S867–S868.