
NEW SOUTH WALES COUNCIL FOR CIVIL LIBERTIES

**SUBMISSION TO THE AUSTRALIAN BUREAU OF STATISTICS
IN RELATION TO PROPOSAL TO DEVELOP
A STATISTICAL LONGITUDINAL CENSUS DATASET**

DATE: 10 June 2005
AUTHOR: Pauline Wright
Vice-President
NSW Council for Civil Liberties

Question 1 – In which state or territory are you/your organisation based?

NSW

Question 2 – Are you completing this questionnaire as a private individual or as a representative of an organisation or government department?

Representative of an organisation or government department.

Question 3 – Which sector best describes your organisation?

Non-Government Organisation.

VIEWS ON THE PROPOSAL TO ENHANCE THE 2006 POPULATION CENSUS

Proposal 1

The ABS is proposing to create the Statistical Longitudinal Census Dataset (SLCD) by:

- ***combining data from the 2006 Census with data from future censuses; and***
- ***combining data from the 2001 Census with data from the 2006 Census***

Data would be brought together without using name and address, using statistical techniques. Name and address would be destroyed after census processing. The SLCD would be for statistical use only and no identifiable data would be released from the ABS.

The reasons for the ABS considering this proposal are explained in the Discussion Paper in Section 2.

Question 4 - Do you, or does your organisation, support Proposal 1 if 2006 Census data are combined with data from future censuses?

No.

Question 5 – What are your reasons?

The proposal outlined in the Discussion Paper represents a radical shift in the nature of the Census. Over the years the ABS has developed a reliable set of statistics because the public has had confidence in it. This proposal because of its significant privacy impacts will erode public confidence in the Census process which will in turn erode the reliability of the data collected.

The proposal will result in an inherent increase in the identifiability of the data and the identifiability will increase over time as the dataset gets deeper. The main issue is "future-proofing": the impossibility of guaranteeing any change in use of information gathered over successive Censuses by government in the future.

The NSW Council for Civil Liberties (CCL) opposes the proposal on the basis that this inherent increase in the identifiability of the data over time will make it tempting in the future for legislation to be changed to allow additional uses not currently contemplated. It is impossible to legislate against this.

The Discussion Paper at page viii indicates that "names and addresses will continue to be destroyed following processing of the 2006 Census", however, there is no set time limit for how long such "processing" may take. Further there is no limit as to who may be allowed access to the data in the future and use of name and address information for specific research projects (see Discussion Paper page 23).

At page 4, the Discussion Paper says “the ABS is proposing to bring together the data from successive censuses through statistical techniques, rather than matching based on name and address”, however the proposed “probabilistic record linkage” techniques will make it relatively simple to re-identify individuals, especially over time.

This involves bringing together datasets by using characteristics such as date of birth, sex, geographic region and country of birth. It should be noted that religious affiliation is another characteristic that could be used for data-matching (Discussion Paper p6). One can anticipate that it would not be difficult to trace the name and address of a (hypothetical) Muslim man born on 10 June 1956 in Afghanistan who came to Australia in 2003 and is living on the Central Coast of NSW.

The Discussion Paper acknowledges that such linkages would for some individuals “be correct” (p6). But even with a degree of statistical uncertainty, over time with a narrow field of possible people matching the description, it would be relatively simple to track an individual if authorities were so-minded.

It is asserted that the SLCD “would be for statistical purposes only and not for administrative or client management purposes” and we acknowledge that this may be the present intention. The difficulty we perceive is that the stated purpose may change over time and the dataset may become subject to use for different purposes via legislative or policy change in the future.

Question 6 – Do you, or does your organisation, support Proposal 1 if 2001 Census data are combined with data from the 2006 Census?

No.

Question 7 – What are your reasons?

In 2001, for the first time, Australians participating in the Census were given the option of enabling their identifying information to be archived. Individuals who agreed to this (and it should be noted that nearly half did not agree, even in the absence of a public campaign against it) may not have done so if they had been made aware that their data was to be made available to develop a longitudinal dataset.

The retrospective merger of the 2001 Census data with future data would deepen the consolidated data with information collected without the consent of the participants for it to be used in that way. People (irrespective of whether they agreed to archiving) participating in the 2001 Census were not given the opportunity to consent to the proposed use of the data in the way proposed. Indeed, the use of this information in the manner suggested may be unlawful.

Proposal 2

The ABS is proposing to make the SLCD available for use in conjunction with other ABS statistical collections, such as ABS household surveys, using statistical techniques. Use would be for statistical purposes only, with no identifiable data released outside the ABS.

The reasons for the ABS considering this proposal are explained in the Discussion Paper in Section 3.

Question 8 – Do you, or does your organisation, support proposal 2?

No.

Question 9 – What are your reasons?

The merging of ABS datasets means enrichment of the longitudinal dataset. Because the data is continually added to and enriched over time, the dataset will become a very attractive source of information for other agencies for purposes that are not currently contemplated, such as "national security" and law enforcement. Future legislative change is possible to enable such uses without legal sanction, especially in an environment where the government of the day has control of both houses of parliament and there is no Bill of Rights to offer protection to affected citizens.

At page 9, the Discussion Paper states that "use of ABS surveys in conjunction with the SLCD would be for statistical purposes only", however, those purposes are potentially broad and could involve use of name and address information (see pp 15 and 23 for discussion of "specified statistical studies" in "exceptional cases"). Further, stated purposes are subject to "functional creep" over time. Powerful authorities may wish to acquire information from the dataset in the "public interest". This is more likely to occur in times where there is a prevailing sense of international insecurity for reasons of national security or law enforcement.

In circumstances where, as now, authorities such as ASIO have been given unprecedented power to detain and question people without arrest or suspicion of having committed any crime for periods of seven days at a time, it would not require any great stretch of the imagination to envisage such authorities being given access to any available information that may be considered by the government of the day to be useful in assisting them with their enquiries.

Proposal 3

The ABS is proposing to make the SLCD available for use, in conjunction with selected non-ABS datasets. These datasets are birth and death register data, long-term immigration data and national disease registers. This would involve using statistical techniques to bring the datasets together. Use would be for statistical purposes only, and no identifiable data would be released from the ABS.

The reasons for the ABS considering this proposal are explained in the Discussion Paper in Chapter 4.

Question 10 – do you, or does your organisation, support Proposal 3?

No.

Question 11 – What are your reasons?

The proposal is for the merger of data acquired by the ABS from censuses and household surveys with data acquired by State and Territory birth and death registries, long-term immigration data (which is defined at page 30 of the Discussion Paper as including “overseas arrivals and departures data, where the period of duration in (sic) over 12 months, and visa grant data, including type of visa” – note there is no time limit for the visa duration) and national disease registers.

Our comments in response to Questions 5, 7 and 9 as to the potential impacts of consolidating datasets by merging data from census to census and censuses to household surveys apply to this proposal also. This proposed merger further enriches the data available and will lead to greater ability to identify and track individuals, even without names and addresses (which we note can be used for “specified statistical projects” – see pp 15 & 23 Discussion Paper).

The Discussion Paper outlines the potential uses for such consolidated data at page 11, including that it could “help inform decisions on the numbers of people brought to Australia..., where they might be encouraged to settle, the amount of assistance they require and the level and location of services required to support these people”. It could equally be used to identify and track individuals of specific nationality and religious affiliation for reasons, for instance, of “national security”. This is something we consider most Australians would find alarming.

Further, the presently intended purposes and limits on the use of the consolidated datasets can be changed over time by legislative amendment. These intended protections cannot be “future-proofed”.

We note further that at page 12 of the Discussion Paper it is “proposed that the quality of the linkage achieved through the statistical techniques be evaluated against a name and address based linkage undertaken by the ABS at the time of census processing”. Clearly then, the retention of name and

address information will be for a significant time after receipt of the census. This is a radical departure from the present situation where name and address information is immediately destroyed upon receipt of census data.

The merging of demographic identifying data coupled with the indefinite retention of name and address data during “census processing” enabling ready accessibility of personal data, presents a real and alarming threat to the privacy and security of Australian citizens and residents. This presents an unacceptable temptation (in the interests of law enforcement and combating perceived threats to national security) for function creep.

Proposal 4

(a) The ABS is proposing to use name and address information, available at the time of census processing, to bring together the 2006 Population Census data with the ABS' 2006 Agricultural Census.

Name and address information used in combining the data would be destroyed on completion of census processing. The dataset would be used for statistical purposes only, and no identifiable data would be released from the ABS.

The reasons for the ABS considering this proposal are explained in the Discussion Paper in Section 5.

Question 12 – Do you, or does your organisation, support proposal 4(a) if the ABS used 2006 Population Census data with 2006 Agricultural Census data?

No

Question 13 – What are your reasons?

See reasons set out above in response to Questions 7, 9 & 11 with respect to merging the SLCD with ABS household survey data and data gathered by other registries.

Residential address would be the key item that would be used to bring together data from the Population Census and the Agricultural Census (p15 Discussion Paper), which of course means that such information would have to be retained during the period of census processing, which has no defined time limit. Accordingly, the merged data will be identifiable back to individuals for a potentially unlimited time.

(b) For quality studies and other particular statistical studies, the ABS proposes to consider the use of name and address information available at the time of census processing to bring together 2006 Population Census data with other selected datasets, namely, ABS household surveys, birth and death register data, long-term immigration data and national disease register data.

Any datasets created for these statistical studies would be kept for a limited time and destroyed after completion of each project.

Census name and address information would be destroyed on completion of census processing. No identifiable data would be released from the ABS.

The reasons the ABS is considering this proposal are explained in the Discussion Paper in Section 5.

Question 14 – Do you, or does your organisation, support Proposal 4(b) if the ABS used 2006 Population Census data with these additional datasets?

No.

Question 15 – What are your reasons?

The retention of name and address data for any time longer than is currently the case is likely to rapidly erode public confidence in the ABS census process and is therefore likely to erode the reliability of the data as individuals are less likely to give complete and accurate information if they have fears that they could be readily identified. This may be especially so for people who have migrated from politically dangerous or oppressive countries where their levels of trust in government or government-funded bodies is limited at best.

It is proposed to destroy name and address information upon completion of “census processing”. There is, however, no time limit set for census processing. It will include the time taken to consider and undertake any specific statistical studies referred to at pp 15 & 23 of the Discussion Paper and to cross-check the quality of data when merging with other datasets (see p12) as well as to analyse the Agricultural Census data. Accordingly, name and address data could be retained for significant and potentially unlimited periods of time.

It is stated at page 14 of the Discussion Paper that the “use of name and address information would be for internal ABS purposes only”, however, it is impossible to guard against future legislative amendment enabling such information becoming available to other agencies not presently contemplated.

“Specified statistical studies” are referred to at page 15 (and again at page 23) of the Discussion Paper. These are studies where “the dataset involved does not have sufficient information for statistical techniques to produce a combined dataset of adequate quality for a significant statistical purpose” and it is proposed that name and address information could be used to bring the dataset together with the SLCD for “specific approved statistical studies” which would “address specific significant research questions and result in new statistical outputs”.

The NSW Council for Civil Liberties opposes the retention and use of name and address data for such ill-defined purposes for potentially unlimited periods of time. This proposal is inherently dangerous and could encompass research for any project the government of the day saw as “significant”. It could include, for instance, research into links between ethnicity, religion and threats to national security and could accordingly pose a threat to the right of individuals to practice their chosen religion.

Privacy and Confidentiality

The ABS abides by a range of privacy and confidentiality provisions. These are discussed in Section 6 of the Discussion Paper.

Question 16 – Are there any additional procedures or protocols you would like the ABS to consider?

The NSW Council for Civil Liberties opposes the proposal, because we do not consider that any additional procedures or protocols could be implemented which could permanently protect the privacy of individuals (because of the possibility of legislative or policy amendments) or prevent abuse of the data or function creep. This is more markedly so because Australia does not have a Human Rights Act or Bill of Rights which might otherwise afford some protections.

Possible Protective Measures

If the proposal as outlined in the Discussion Paper were to be introduced notwithstanding opposition to it, the minimum necessary safeguards would include the following:

- (a) legislative change to create additional offences for misuse of census information and/or increase penalties;
- (b) provide private rights of action in the event of misuse, including the ability to obtain injunctive relief, and compensation for people adversely affected;
- (c) ensuring all uses of longitudinal datasets are subject to ethics committee reviews, and that the committees have majority community representation;
- (d) adequate disclosure of potential uses of census data at time of collection;
- (e) opt-in or opt-out rights;
- (f) destruction of longitudinal datasets in times of war.

A failure to adopt the above will lead to significant reduction in public confidence in the census, which will probably increase over time as the data becomes more specific. By then it will be too late to restore public confidence.

One possibility for improvement of the proposal would be to base the longitudinal dataset on data other than name and address which could produce a statistically sufficient degree of certainty, but sufficiently less than 100% so that it is not reliable enough for anyone wanting to use it for a non-statistical purpose.

Another possibility is that the SLCD be restricted to 10% of the total census population. As we understand it, the UK have a SLCD based on a 1% sample.

The absence of a Human Rights Act or Bill of Rights in Australia, providing individuals with redress in the case of abuses, poses a major impediment to any impingement on the anonymity of census data being acceptable.

Current and Proposed Protections

The provisions of the Privacy Act 1988 and the Census and Statistics Act 1905 afford some protection at present because they impose criminal sanctions for abuse of personal information. Such legislative protections are only good, however, while they are in force and while they remain unamended. In circumstances where the government of the day has control of both houses of Parliament, it would not be a difficult matter to amend the protective legislation to enable agencies other than the ABS to have access to the SLCD. It should also be noted that there are many exceptions to general limitations on the use of personal information contained in the present legislation.

The Privacy Impact Assessment is a laudable initiative, but is likely to reach conclusions, we believe, similar to those which we have reached in relation to the privacy impacts and impossibility of “future-proofing” the SLCD. Furthermore, there is no requirement to abide by the recommendations of the PIA.

We note further that ABS officers are bound by secrecy provisions and undertakings and are subject to sanctions if they breach those provisions and undertakings. Nevertheless there are always unscrupulous individuals willing to risk possible sanction for certain rewards.

It is also proposed that the initiative be overseen by the ASAC. While the individuals on the ASAC are appointed by government, public confidence of their independence may not be high, especially in the face of such potentially far-reaching privacy impacts as are posed by this proposal.

Why Protective Measures Will be Inadequate

By way of summary, we consider that no procedure or protocol can provide adequate protection because:

- a. national security and law enforcement agencies can be powerful lobbyists and it is inevitable that where such a rich source of data is available, other agencies will seek access to it;
- b. ABS policy is subject to change;
- c. legislation can be changed and that will be made more simple once the government has control over both houses;
- d. it may be possible for government agencies to provide de-identified information to the ABS (which could then be re-identified using probabilistic techniques) to ensure that they do not run foul of the present legislation;
- e. functional creep is possible and, indeed, likely (consider for example the use of TFN for data-matching very soon after it was introduced notwithstanding the fact that we were assured that it would not be so used);

- f. ASAC will not have public confidence because of perception of lack of independence given that it is government-appointed.

Question 17 – If the ABS adopted any of these proposals, would you, or your organisation, be likely to apply to use the SLCD, either in its own right or conjunction with one of the other datasets specified in the Discussion Paper?

No.

Question 18 – If yes, what types of projects would you use the datasets for, why would you need the SLCD (or a combination of the SLCD and other datasets) and if known, which data items would you need?

Not applicable.

General Comments

Do you have any other comments on the proposal to enhance the population census dataset?

The Discussion Paper at p1 rightly identifies the fact that the success of the ABS to date “is built on the high level of community support for, and trust in, the ABS, which in turn is based on the sound reputation it has developed for consistently placing high priority on managing the privacy of information entrusted to it.” These proposals will bring that sound reputation into jeopardy and will risk the long-term reliability of the data as community trust diminishes.

The ABS has the power to compel citizens to take part in the census. This power has historically been used responsibly and with great sensitivity for the personal nature of the information being gathered. This proposal is a radical shift from that position and poses a real threat to the privacy and security of Australians. The NSW Council for Civil Liberties believes that if adopted, the proposals will lead to a significant reduction in public confidence in the census, which will probably increase in time as the dataset becomes deeper. By then it will be too late to restore public confidence.

The proposals outlined in the Discussion Paper would mean the creation of a national population register or database. In the mid-1980's when the Australia Card proposed a national database, it was soundly rejected by the Australian people. The longitudinal dataset being now under consideration would be a much deeper dataset than that which was proposed under the Australia Card. If the general public understood this, the outcry will be deafening.

It should not be forgotten that the Nazi regime and the consequent retreat from civilization experienced in Germany in the lead-up to and duration of World War II came about from a properly-constituted democratically-elected government. Access to a national population dataset of the nature contemplated would have made the task of locating and detaining and the ultimate genocide of Jewish people much simpler for that regime.

The absence of a Human Rights Act or Bill of Rights in Australia providing individuals with redress in the case of abuses is a major impediment to agreeing to any impingement on the anonymity of census data being acceptable.

What next?

Information from this submission will be used to inform the Australian Statistician about the reaction of the Australian public, census data users and other interested people, to the proposal being considered. An 'in principle' decision on whether to proceed with any part of the proposal will be made by late June 2005. A final decision will depend on community support and the legislative process. You will be notified of the decision.