

Submission to the Australian Bureau of Statistics
on its proposal –



Enhancing the Population Census:

Developing a Longitudinal View

June 2005



Office of the
Victorian Privacy
Commissioner

TABLE OF CONTENTS

1	Introduction.....	1
2	Will the proposal take Australia too close to the creation of a universal and unique identifier?	1
3	Distinguishing different categories of data in order to consider the SLCD proposal more clearly.....	2
	Data to be collected	2
	Data to be retained in irretrievably de-identified form because names and addresses have been removed	2
	Data to be retained but in a form that means it can be re-identified.....	3
	Data to be published.....	3
	Data that is re-identifiable and is to be made conditionally available in deidentified form to approved recipients for specific purposes	4
4	Does the ABS have the necessary authority to collect, use and disclose data for longitudinal research?	4
5	Will the creation of a SLCD be overly intrusive?	5
6	Can the ABS sustain its pledge of confidentiality in the face of compulsory powers?	5
7	In a census that will produce a SLCD, what would be a “reasonable excuse” for refusing to provide data?.....	6
8	How might the quality of census data suffer?	7
9	Why can’t the SLCD use the responses of only those people who give informed consent?	8
10	The importance of trust.....	9
11	Privacy Impact Assessment	11
	Endnotes	12

The Privacy Commissioner wishes to acknowledge the work of Michelle Fisher (Manager, Policy, Privacy Victoria) in the preparation of this Submission.

Office of the **Victorian Privacy Commissioner** (Privacy Victoria)
GPO Box 5057
10-16 Queen Street
Melbourne Victoria 3000 Australia
Phone: +61-3-8619-8719
Fax: +61-3-8619-8700
Email: enquiries@privacy.vic.gov.au
Website: www.privacy.vic.gov.au

Submission to the Australian Bureau of Statistics on its proposal for *Enhancing the Population Census: Developing a Longitudinal View*ⁱ

1 Introduction

1. The Australian Bureau of Statistics (ABS) proposes to enhance the value of the 2006 census by linking it to data obtained from future censuses, possibly to the 2001 Census data, to data obtained from other ABS sources such as survey data, and to data sources outside the ABS such as births and deaths data, and disease registers. This linked dataset will be known as the Statistical Longitudinal Census Dataset (SLCD).
2. For brevity's sake, I have not rehearsed the public interests in the research that the Statistical Longitudinal Census Dataset would facilitate. Of course, I acknowledge them. But the public interest in research is always to be weighed with other public interests, and two of them in this context are the public interest in respect for personal privacy and the public interest in maintaining trust in the ABS in its capacity as collector and guardian of census data. Relevant also to this proposal is one of the basic tenets of research ethics: that voluntary informed consent be obtained from research subjects. While not invariable, this tenet is foundational, and it informs my view, explained later, that the SLCD ought not be compiled through compulsion.

2 Will the proposal take Australia too close to the creation of a universal and unique identifier?

3. The ABS states that name and address information will not be retained beyond the time needed to process each census. However, the Discussion Paper makes clear that:
 - a. individuals' information from consecutive censuses or different datasets will nevertheless be linked by using other identifiers, such as age/date of birth, sex, geographic region, and country of birth (page 6);
 - b. for most cases, the linked datasets would reflect the data of the same individual (see page 24), although the linkage may not be certain in all cases (page 6); and
 - c. the quality of the linkage would be tested against a name and address based linkage during the time of census processing (at page 12).
4. Whether or not names or other identifiers are used, the aim of the SLCD is the same: to create a longitudinal view of individuals as they move through transitions during their lives (page 1).
5. The process of bringing different datasets together and associating them with individuals will be achieved through the use of a key, which is based on the above identifiers (pages 6 & 19). This key will allow different records to be linked to the same individual, with a high degree of probability. As the census is universal in its scope – involving the collection and retention of data from every citizen, resident and visitor in Australia on census night (excluding foreign diplomats and their families) – it

appears that the ABS will be creating a universal and (relatively) unique identifier. Serious privacy concerns are raised by any initiative that involves the creation of any universal identifier. The “Australia Card debate” of the 1980s demonstrated those concerns, and technological developments since that time have not diminished, but rather have increased, the potential for unacceptable intrusions into privacy that a universal identifier (combined with census responses) would enable.

6. In my opinion, if such a universal identifier is ever to be created, it should be only by a considered express decision of the Parliament, taken after due process and a proper opportunity for wide and informed public debate. While the efforts by ABS to consult the public on this proposal are acknowledged, they are not commensurate with the significance of what is proposed, nor likely to obtain the attention that action by the Commonwealth Parliament would attract.
7. If a universal identifier and SLCD were to be created under purpose-built law, separate issues would arise about the clarity and limits of authorised uses and about safeguards against misuse of such a powerful dataset of the population. The SLCD material can be expected to be highly attractive to a range of government agencies (eg Australian Taxation Office, Centrelink, Department of Immigration and Multicultural and Indigenous Affairs, Department of Foreign Affairs and Trade, and state and federal security and law enforcement agencies), as well as private sector entities specialising in data mining and profiling.

3 Distinguishing different categories of data in order to consider the SLCD proposal more clearly

8. Consideration of the SLCD proposal will be clearer if categories of data are clearly distinguished. The categories matter because discussion of the proposal can become confused if assurances that apply to some categories are misunderstood to be capable of applying to all.

Data to be collected

9. This means answers to questions to be asked in the census, most of which are compulsory, obtained in identified form at the time of collection and during enumeration of the census.
10. This also includes the other datasets which, together with the census, will comprise the SLCD. When those other datasets were compiled, certain assurances will have been given. The use of those datasets in the SLCD may undermine those assurances.

Data to be retained in irretrievably de-identified form because names and addresses have been removed

11. It is popularly believed that census data is irretrievably de-identified. Most people understand that, with the exception of those who consented in 2001 to the retention by the Australian Archives of their data in identifiable form for disclosure in 99 years, the ABS does not keep links between the data and the people who supplied it after the ABS has finished processing the census forms because the ABS destroys names and

addresses. If this popular understanding is not accurate, it would be in the public interest for the ABS to explain precisely how it treats identifiable census data.

12. Removal of names and addresses would not appear to achieve irretrievable identification. If done properly, de-identification removes privacy issues. But proper de-identification is complex, especially in an era of powerful information processing technologies.
13. Removal of names and addresses, but retention of codes may allow individual census responses to be linked back to that individual, although only the ABS may know what the codes mean, and only the ABS may be able to make the link.

Data to be retained but in a form that means it can be re-identified

14. This category will often be “personal information” within the meaning of the privacy laws. Even shorn of name and address, data can be re-identified. Whether identity is reasonably ascertainable will depend on the particular circumstances, including the datasets that are available for matching and the precise forms of any coding that may have been used. It will often be relevant to examine the resources and purposes of the entity that may collect the data and attempt the re-identification.
15. The re-identification may be possible for ABS using its own internal codes, or for another organisation with capacity to link and match other datasets.

Data to be published

16. It is a basic routine function of ABS to publish de-identified data. This data does not raise privacy issues when it is aggregated for the whole nation or for large geographic areas (macro data) and it is not possible to identify any of the individuals whose personal information helps to form the data.
17. The ABS also provides access to unit record data (or micro data) in the form of “Confidentialised Unit Record Files” (CURFs). According to the ABS, CURFs “contain the most detailed statistical information available from the ABS, and are of most use to researchers and analysts who wish to run their own statistical queries on the data” and are ‘confidentialised’ by:
 - removing name, address and any other information that might uniquely identify an individual;
 - changing a small number of values, particularly unusual values, and removing very unusual records
 - controlling the detail available for all records on the CURF
 - controlling the modes of access to restrict access to more detailed data
 - placing restrictions on how the data are used, supported by a legal undertaking signed by each user and their organisationⁱⁱ
18. Privacy issues can arise when the geographic area from which the aggregated data is formed is sufficiently small, or sparsely populated (even if geographically large), or contains details that appear so rarely (alone or in combination with each other) that

identification of an individual becomes possible (meso data). If the particular detail includes, say, ethnicity or wealth or health, then at certain times some members of the community can be made vulnerable by being able to be re-identified.

19. I acknowledge that ABS is acutely conscious of de-identification issues and takes pains to ensure that the data it publishes is appropriately de-identified. It should be apparent how significant it is to handle appropriately both micro data and meso data.

Data that is re-identifiable and is to be made conditionally available in de-identified form to approved recipients for specific purposes

20. This is the SLCD. The purpose is to allow research over time into aspects of the lives of individuals in order to study their various life experiences – say, unemployment, single parenthood, or results of exposure to a health hazard – and make responses based on what can be learned. The proposal cannot work unless people are able to be re-identified with reasonable precision to enable their disparate records to be linked so as to enrich what can be known about them over time (longitudinally).
21. It is not suggested that this re-identifiable data is proposed to be made public by the ABS – or any of the recipients to whom ABS provides it – in identifiable form. Assurances that the ABS will not do this, assurances that ABS will bind recipients to conditions that aim to prevent this, and the assurances that ABS will keep secure in a special unit of ABS the linkage keys that will be able to link an individual back to his or her data, are all accepted in good faith, but they do not deal with the privacy problems inherent in the SLCD.
22. The privacy problems stem from more fundamental concerns about the compulsory, universal collection of data that many people will regard as unreasonably invasive if it is to be retained in re-identifiable form and given to an unknown range of organisations for purposes yet to be definitively explained under conditions that may not be able to be enforceable in future.

4 Does the ABS have the necessary authority to collect, use and disclose data for longitudinal research?

23. The ABS is given wide powers under the *Australian Bureau of Statistics Act 1975 (Cth)* to collect, compile, analyse and disseminate statistical information, and to maximise the utilisation of information obtained from official bodies for statistical purposes. The ABS is also given authority under the *Census and Statistics Act 1905 (Cth)* to conduct the 5-yearly census and to collect data on the number of people in each State and on any matter the Statistician considers appropriate. Census information is relied on for a number of purposes, including estimating the population of the States and Territories, determining representation entitlements under the *Commonwealth Electoral Act 1918*, distributing Commonwealth Grants to the States and Territories, and for forecasting and policy making by government. None of these purposes requires that identifiable data be retained.
24. It is not clear whether the relevant laws provide the ABS with sufficient power or authority to implement the SLCD proposal to bring together the census with other datasets, link them over time, retain them in (re)identifiable form, and make them

available to researchers and others interested in how various factors influence individuals' experiences over time.

5 Will the creation of a SLCD be overly intrusive?

25. Under the census law, the Statistician can ask anything he considers appropriate to ask Australians, and respondents must answer on pain of a continuing daily fine.ⁱⁱⁱ Before each census is taken, census topics (but not the specific questions) must be set out in regulations and tabled in Parliament, and this is a disallowable instrument. In such circumstances, what is it appropriate to ask Australians? The consensus view may change if the traditional assurance of anonymity is to be replaced by an ongoing intention to retain responses in identifiable form and to link them over time to other data that is also able to be linked to the person responding.
26. The census and ABS surveys contain a number of questions that individuals may find overly intrusive and invasive, especially if their answers are kept in re-identifiable form and combined with other data over time. For instance, it is proposed that the 2006 Census will ask individuals about their religion, their and their parent's ethnicity, their Indigenous status, their income level, their family relationships (potentially including same sex de facto relationships), and number of children ever born (potentially including stillborn babies, or babies born in delicate circumstances not necessarily known to others in the person's current relationships). Household Expenditure Surveys require individuals to inform the ABS of their financial details, including their income and every purchase over a defined period of time, and indicators of deprivation or financial stress. The ABS also collects health information through the National Health Survey. The Agricultural Census frequently produces an overlay of family affairs with the business affairs of a farm or station, as the Discussion Paper notes at pages 14-15.
27. Sensitive questions, such as those relating to racial and ethnic background, should be optional. If a SLCD were to be created in future, that dataset should not contain racial or ethnic data. These are the ingredients of meso data that can produce serious problems.

6 Can the ABS sustain its pledge of confidentiality in the face of compulsory powers?

28. It is often said that concerns about privacy and the census can be allayed by the legislative safeguards already in the *Census and Statistics Act* and the *Privacy Act*. These assurances are fragile now and not immune from removal or change in the future. The current fragility of *Census and Statistics Act* safeguards stems from the fact that they may give way to an order to compel disclosure.^{iv} With the limited exception of section 19A of the *Census and Statistics Act* (relating to non-disclosure of 2001 Census data in court proceedings), the ABS appears to have no legislative authority to resist a demand – whether by warrant, subpoena or compulsory power exercised by a Commonwealth agency – for the information it holds. A standard feature of privacy laws is that their protections give way when disclosure is required or authorised under another law, and in several other contexts that may have application to census data.^v

29. As to the future, it may be argued of course that the Statistician cannot be held accountable for what others might do. But it is possible in the current environment to take account of several trends that may lead to legislative change that would trump census-related assurances.^{vi} This could only occur if the data were re-identifiable.
30. As the saying has it, ‘what can be known, will be known’. To make it impossible to re-identify census data is, in effect, to take the pragmatic step of ensuring that what cannot be known will not be known. If the census data, compulsorily acquired, is retained in re-identifiable form, it will be possible for it to be put to other uses under laws passed in future. If the ABS makes assurances that prove to be fragile in practice, its reputation will suffer in the long run (see Section 10 later). In this specific context, I note that the Discussion Paper raises the prospect of linking 2001 Census data into the SLCD, and this would appear to run counter to assurances given to the nation when it completed its forms in 2001.
31. The assurances contained in State and Territory laws in relation to datasets that the ABS proposes to incorporate into the SLCD are unlikely to operate to protect the privacy of those who supplied the data to those State and Territory entities (eg births, deaths and marriages registers). This is because these laws would give way to a Commonwealth law that is inconsistent with them (Commonwealth Constitution, s. 109).
32. In light of the above, and while I accept the good intentions of the ABS, it does not appear that true “functional separation” can be guaranteed in the way the Discussion Paper suggests at page 19.

7 In a census that will produce a SLCD, what would be a “reasonable excuse” for refusing to provide data?

33. Section 14(1) of the *Census and Statistics Act* provides that a person may refuse or fail to comply with a direction [to provide information] with reasonable excuse. Consistent with general principle, preservation of rights and liberties (eg against self-incrimination) would appear to be a reasonable excuse in a census context. Re-identifiable data will be of interest to government entities with an enforcement role. Although safeguards are suggested, for the sake of this discussion, let it be assumed that “function creep” – a phenomenon commonly encountered in privacy circles and relevant, for example, in the history of the Tax File Number – is likely to occur to some extent. This is not a novel argument. It has been articulated by others:

The present practice of destroying identified census information as soon as it has been processed decreases the chances of such an issue [compulsory disclosure] arising in practice. The information supplied in a census form may reveal infringements against taxation and social security laws, particularly if questions are asked about relationships and income. If it were possible for that information to be made available to courts and to other bodies with powers to make decisions adverse to the interests of a census subject, two main objections would arise. First, privacy interests suggest that members of the public should be warned in advance of this possibility. Such a warning might well affect the validity of the information supplied by some persons in response to certain questions. Secondly, it might be thought wrong in principle to require that, in answering census questions, a person incriminate himself in respect of offences committed by him.^{vii}

8 How might the quality of census data suffer?

34. The paragraph above also contains an implicit warning of risks to the quality of census data if the SLCD proposal is adopted. Where people believe that questions are overly intrusive and/or that safeguards are not adequate, the quality of their answers may be affected. This is an important factor because it is in the public interest that the quality of census data be high in order for the census to serve its main public interest purposes. Those purposes go to the effectiveness of an elective democracy, the distribution of revenue in a federation, and the quality of decision-making in many diverse areas of public policy and commerce.

35. The SLCD proposal seems to indicate a change in longstanding ABS policy. During a 1997 Parliamentary inquiry into the treatment of census forms, the Australian Bureau of Statistics referred to their longstanding support for the destruction of census forms:

All Australian Statisticians since 1971 have strongly supported the policy on the destruction of census forms.

In an address to the Australian Archives Council in November 1995... the current Australian Statistician, Mr W. McLennan, set out his views on the current policy. Mr McLennan expressed serious concerns that a possible adverse public reaction to the retention of name-identified census records would affect public confidence and trust not only in the census, but in other data collected by the ABS, and thereby affect public cooperation in the collection of official statistics. This would affect the quality not only of census data, but also of data from other ABS collections. A reduction in data quality from the census would diminish its value for these important purposes and possibly put some of them at risk.^{viii}

36. It is in the public interest for everyone to respond to the census. In Australia, that includes groups that may have legitimate reasons to be wary of compulsory information collection that seems overly intrusive and will remain identifiable. For example, the large proportion of Australians who have immigrated from countries where government use of their personal information could be far from benign. The wary also may include Indigenous Australians in whose family memory is the experience of the Stolen Generation.^{ix}

37. One writer has emphasised the link between confidence and quality as follows:

Even when responses to requests for information are required by law, the success of a statistical programme depends in large measure on the willing cooperation of respondents. Respondents who understand the purposes of the inquiry, who sympathize with the intended uses of the information, and who believe that providing the government with the requested information will not harm them are much more likely to answer truthfully and with a minimum of effort on the part of the data-collection agency.^x

38. One of the most prominent scholars in this field has distilled from a disparate set of examples the following table of factors in population statistics work that, unless carefully considered, may lead to potential harm. Longitudinal data gathering is mentioned:

Table 3. Factors Contributing to Higher Risk of Population Data Collection Effort based on Potential for Respondent or Group Harm^{xi}

A. Critical factors
<ol style="list-style-type: none"> 1. Population studied is weak or otherwise vulnerable. 2. Data gathering or research involves variables that are on _ sensitive _ topics, typically topics that are or can be used to identify or stigmatize one or more vulnerable groups, or use classifications that permit the identification or stigmatization of such groups.
B. Aggravating factors
<ol style="list-style-type: none"> 1. All or substantially all of population is covered, i.e., sampling is not used. 2. Longitudinal data gathering is involved, or the activity can be linked to a longitudinal system. 3. Participation is mandatory or is effectively coerced. 4. Little or no input from the subject population in planning the data gathering or research activities. (The risk potential is further enhanced if there are substantial inputs in terms of expertise, staff, or funds from foreign persons or institutions.) 5. The data gathering or research is carried out in a war, a period of civil disruption, or during or shortly after a similar emergency. 6. Little or no attention given to organizational, operational, methodological, and technological safeguards against the misuse of information obtained for non-statistical purposes. 7. Confidentiality assurances provided to respondents have limited or no legal basis. 8. Ethical reviews are not carried out, are perfunctory, or are heavily influenced by utilitarian considerations.

Note: The presence of either or both critical factors gives rise to a presumption of risk and each additional aggravating factor present further augments such risk. On the other hand, it should be emphasized that the presence of critical and aggravating factors does not mean that actual harm has occurred.

9 Why can't the SLCD use the responses of only those people who give informed consent?

39. Informed consent is a touchstone in privacy and in research. The census is the “mother of all surveys”. It covers about 22 million people. It would appear that, if respondents were asked to consent to the retention in re-identifiable form of their personal information for research purposes with safeguards, a statistically meaningful number would agree. In the 2001 census, approximately half of all census subjects consented to the retention by the Australian Archives of their personal information in identifiable form for disclosure in 99 years. To the non-expert, 11 million or so looks like a pretty good sample. The ABS needs to explain to non-experts why it isn't.
40. Given all the factors to be weighed in this matter of a SLCD, the onus is on the ABS to demonstrate that the overwhelming public interest justifies the use of compulsion to collect and keep in re-identifiable form for future uses not comprehensively listed such

a range of sensitive data. The extent of the collection seems disproportionate, especially when weighed against the concerns listed above and below. In my opinion, the ABS is yet to discharge this onus.

10 The importance of trust

41. The brittleness of public confidence in a population-wide census is international, and is justified by some painful historical lessons. These are to be found in a literature too large to be cited other than briefly here. That literature, which will be well known to specialists in the ABS, deals with the role of national statistical agencies and the data they collected in episodes including:
 - a. misuse of ID card and census data in Nazi-occupied countries to facilitate the Holocaust, and the long-term impact on the collection of population statistics in the affected countries (- the extent to which various national statistical agencies were knowingly complicit appears to be a matter of dispute among some);^{xii}
 - b. assistance given by US census authorities to the US War Department to facilitate its round-up of Japanese Americans for internment after the attack on Pearl Harbour in December 1941 (- the extent of that assistance is in dispute among experts);^{xiii}
 - c. administration of the official policy of apartheid in South Africa (from the early 1950-1993);^{xiv}
 - d. targeting Tutsis in genocide in Rwanda in 1994 (a scholar's tentative finding, pending further research);^{xv}
 - e. implementation of policies now understood to have contributed to the harms done to indigenous peoples in Norway, Australia and the United States.^{xvi}
42. Other indications of public wariness can be found. For example, the UK poll tax debate harmed the response rate in the 1991 UK census, the results of which were so poor they were not used as the basis of population estimates and the 1981 census result continued to be used.^{xvii} Consider also the periodic expressions in Australia of public disquiet (eg 1971 and 1976) about the compulsory nature of the census, or about the questions it asks.^{xviii} Among the responses to this disquiet was a promise of permanent de-identification of census responses.^{xix} Consistent research findings for the ABS.^{xx} in relation to confidentiality of personal information, including census data, indicates that people are concerned about it.
43. Rather than try to weigh here the contending views about the extent to which various national statistical agencies may have assisted the historical episodes listed above, I draw the Statistician's attention to a more subtle message from each of these episodes. The unmistakable message is in two parts:
 - a. if a nation's statistical agency retains data that has a reasonable chance of being used to re-identify persons in particular groups with particular characteristics, then that agency may be requested or required to provide the data; and
 - b. if that data, which may have been compulsorily collected by the agency, is used in ways that do harm to members of a particular group, then it is that agency that will

be dogged for years and years by the episode (regardless of what happens to the government responsible for the harms).

44. The literature referred to above contains continuing debates in which the relevant national statistical agencies are necessarily embroiled. The historical episodes – whether proven, half-right, or unproven - resurface in the media.
45. Just recently, when the US Census Bureau provided data to the Department of Homeland Security about Arab Americans (data that is public and that the Department could have obtained itself from the Bureau’s website), the *New York Times* reported the matter alongside a reference to the still-disputed role of the Bureau in the round-up of Japanese Americans after Pearl Harbour in 1941.^{xxi} These matters require of the statistical agency fresh repeats of earlier public statements about the historical matter – and these might be repeats of regrets, or of corrections to misunderstandings, or of denials. One illustration will suffice. In 2000, the then Director of the US Census Bureau wrote in an email to staff:

The historical record is clear that senior Census Bureau staff proactively cooperated with the internment, and that census tabulations were directly implicated in the denial of civil rights to citizens of the United States who happened also to be of Japanese ancestry.

The record is less clear whether the then in effect legal prohibitions against revealing individual data records were violated. On this question, the judicial principle of innocent until proven otherwise should be honored. However, even were it to be conclusively documented that no such violation did occur, this would not and could not excuse the abuse of human rights that resulted from the rapid provision of tabulations designed to identify where Japanese Americans lived and therefore to facilitate and accelerate the forced relocation and denial of civil rights.

I would also like to state clearly that for many years the Census Bureau was less than forthcoming in publicly acknowledging its role in the internment process. Silence was not the worst offense, for there is ample evidence that at various times the Census Bureau has described its role in such manner as to obfuscate its role in internment. Worst yet, some Census Bureau documents would lead the reader to believe that the Census Bureau behaved in a manner as to have actually protected the civil rights of Japanese Americans. This distortion of the historical record is being corrected.^{xxii}

46. In 2005, the US Census Bureau is still dealing with the consequences of 1942:

Last year, a *New York Times* article regarding tabulations the U.S. Census Bureau provided to another agency, the U.S. Department of Homeland Security, highlighted certain statements and allegations. The article drew comparisons to actions in the Census Bureau’s past, and suggested the Census Bureau—by providing statistical data, which did not reveal any individuals—had acted in bad faith, violating the trust under which it collected the census from individuals and households. The article compared the preparation of extracts for the U.S. Department of Homeland Security to the work that the Census Bureau did in 1942 for the War Department, specifically the Army’s Western Defense Command in San Francisco, during World War II, for the purposes of relocating the Japanese community living in the west coast. The comparison between the tabulations the Census Bureau provided to the Department of Homeland Security and the work it did for the army rekindled long-standing doubts and allegations about the Census Bureau’s commitment to *the confidentiality protections of the Census Law*, and *the extent to which statistics are used to the detriment of certain populations*. [original emphasis]

It is important to recognize these are separate issues or concerns, and to confuse and conflate these allegations unnecessarily undermines public trust. However, that the *New York Times* piece so readily rekindled the doubts about 1942, and conflated the issues of law and the use of statistics, caused concern for the Census Bureau. This concern resulted in a review of our responsibilities, the commitments we make to respondents, and our policies.^{xxiii}

47. For present purposes, the point is that it is the statistical agency, not the long-gone government of past days, that continually has to shore up trust so that it can undertake effectively its work in the present and future. It is the statistical agency that continues to have a fundamental interest in maintaining public trust in its integrity in order to maintain the data flows that are critical to its ongoing work.
48. Whatever the truth of these historical debates, the present reality for the statistical agencies involved in them is persistent, debilitating doubt about whether the current leadership can be trusted, and consequently whether the agency's current promises will be kept and whether contemporary safeguards will count for anything in the future.
49. It would appear that a statistical agency pays a certain price in lost trust, regardless of the consultation it undertakes, the safeguards it crafts, or the scrupulousness with which it applies those safeguards.
50. The best response, from the viewpoint of the long-term interests of the statistical agency, would appear to be that the statistical agency should irretrievably de-identify the data that is provided in identifiable form in the census. What can't be known, won't be known.
51. The providers of the personal information (much of it sensitive) are then guaranteed continuing anonymity in the face of, for example, changes to the agency's leadership, lost corporate memory, transformed organisational culture, repealed or amended legal frameworks, shifting public attitudes towards certain minorities, and external pressures created by the transient passions of the times.
52. History contains lessons of relevance to the consideration of the SLCD proposal. On balance, those lessons suggest that the proposal in its present form should not proceed.

11 Privacy Impact Assessment

53. I welcome the Statistician's efforts to generate public debate about the SLCD proposal and to commission specifically for it a Privacy Impact Assessment ("PIA"). I urge the Statistician to publish the results of the PIA and allow time for the PIA to inform the debate.

PAUL CHADWICK
Victorian Privacy Commissioner
9 June 2005

Endnotes

- ⁱ Australian Bureau of Statistics, *Enhancing the Population Census: Developing a Longitudinal View*, discussion paper no. 2060.0, released 26 April 2005, [http://www.ausstats.abs.gov.au/ausstats/free.nsf/Lookup/43B2F651F627AEFACA256FEE00804245/\\$File/20600_2006.pdf](http://www.ausstats.abs.gov.au/ausstats/free.nsf/Lookup/43B2F651F627AEFACA256FEE00804245/$File/20600_2006.pdf).
- ⁱⁱ Australian Bureau of Statistics, *ABS Confidentialised Unit Record Files*, undated, [http://www.abs.gov.au/websitedbs/D3110129.NSF/0/95b232c3f39b2022ca256f4a001098bb/\\$FILE/CURF%20Brochure%20-%20April%202003.pdf](http://www.abs.gov.au/websitedbs/D3110129.NSF/0/95b232c3f39b2022ca256f4a001098bb/$FILE/CURF%20Brochure%20-%20April%202003.pdf).
- ⁱⁱⁱ Under the Census and Statistics Act, a failure or refusal to provide information when requested – \$100 fine (s. 14); providing false or misleading information – \$1000 fine (s. 15). A person may refuse to answer questions or provide information if it concerns their religious beliefs (s. 14(2)), or if they have a reasonable excuse (s. 14(1)).
- ^{iv} See the discussion of the Australian Law Reform Commission’s *Privacy and the Census*, Report No. 12, 1979, para 64, <http://www.austlii.edu.au/au/other/alrc/publications/reports/12/>. Compare with the more explicit disclosure avoidance protections for US census data in section 9 of Title 13 of the US Code (available at <http://uscode.house.gov/search/criteria.php>), the development of which is discussed in US Census Bureau, *Census Confidentiality and Privacy: 1790-2002*, 2002, <http://www.census.gov/prod/2003pubs/conmono2.pdf>.
- ^v See the use-and-disclosure principles of national and state and territory privacy laws, in particular NPP 2.1 (g) of the Commonwealth *Privacy Act 1988*, and Victorian *Information Privacy Act 2000* s. 6 and IPP 2.1 (g).
- ^{vi} The context for consideration of this proposal includes, inter alia, various identity management initiatives at federal and state levels, a Medicare smartcard proposal and a security environment that changed on 11 September 2001, and is different from the environment at the time of the previous census.
- ^{vii} Australian Law Reform Commission, *Privacy and the Census*, Report No. 12, 1979, para 64, <http://www.austlii.edu.au/au/other/alrc/publications/reports/12/>.
- ^{viii} Australian Bureau of Statistics, *Information Paper: House of Representatives Standing Committee on Legal and Constitutional Affairs Inquiry into the Treatment of Census Forms – Submission from the Australian Bureau of Statistics*, 18 July 1997, <http://www.abs.gov.au/websitedbs/D3110122.NSF/0/a956800a923efbdf4a25667300095d05?OpenDocument>.
- ^{ix} A number of culturally and linguistically diverse respondents in a survey conducted on behalf of Privacy Victoria in 2002 objected to being asked questions about race and ethnicity: see, eg, pages 75 (Arabic), 94 (Somali community), 99 (Turkish community), Office of the Victorian Privacy Commissioner, *Privacy in Diverse Victoria: Attitudes towards information privacy among selected Non-English speaking background and Indigenous groups in Victoria*, Report 02.02, October 2002, available at <http://www.privacy.vic.gov.au> > Publications > Reports. See also Ellen Percy Kraly, “Blankets, Brass Tags and Bungalows: The Role of Population Data Systems in Historical Aboriginal Affairs in Northern Territory, Australia”, paper to be presented at the International Union for the Scientific Study of Population, XXV International Population Conference at Tours, France on 18-23 July 2005, <http://iussp2005.princeton.edu/download.aspx?submissionid=50449>.
- ^x Martin, M.E., “Statistical legislation and confidentiality issues” (1974) 42(3) *International Statistical Review* 265–281 at page 265, quoted in George T Duncan et al (eds), *Private Lives and Public Policies: Confidentiality and Accessibility of Government Statistics*, 1993, at pages 28-29, <http://www.nap.edu/books/0309047439/html/index.html>.
- ^{xi} William Seltzer, “Ethics Issues in Data collection”, in *Encyclopedia of Population*, Paul Demeny and Geoffrey McNicoll, eds., 2003, pages 195-197, included at page 57 in William Seltzer and Margo Anderson, “Government Statistics and Individual Safety: Revisiting the Historical Record of Disclosure, Harm, and Risk”, prepared for a presentation at a workshop, Access to Research Data: Assessing Risks and Opportunities,

organized by the Panel on Confidential Data Access for Research Purposes, Committee on National Statistics (CNSTAT), Washington DC, 16-17 October 2003, revised draft for comment 1 December 2003, <http://www.uwm.edu/%7Emargo/govstat/WS-MAcnstat.pdf>, quoted with permission.

xii William Seltzer and Margo Anderson, “Government Statistics and Individual Safety: Revisiting the Historical Record of Disclosure, Harm, and Risk”, prepared for a presentation at a workshop, Access to Research Data: Assessing Risks and Opportunities, organized by the Panel on Confidential Data Access for Research Purposes, Committee on National Statistics (CNSTAT), Washington DC, 16-17 October 2003, revised draft for comment 1 December 2003, pages 32-34, <http://www.uwm.edu/%7Emargo/govstat/WS-MAcnstat.pdf>, quoted with permission.

xiii US Census Bureau, *Census Confidentiality and Privacy: 1790-2002*, 2002, pages 14-17, <http://www.census.gov/prod/2003pubs/conmono2.pdf>. Compare with Hermann Habermann (Deputy Director, US Census Bureau), *Ethics, Confidentiality, and Data Dissemination*, 2005, esp pages 1-7 and 11-14, <http://unstats.un.org/unsd/ethics.pdf>.

xiv William Seltzer and Margo Anderson, “Government Statistics and Individual Safety: Revisiting the Historical Record of Disclosure, Harm, and Risk”, prepared for a presentation at a workshop, Access to Research Data: Assessing Risks and Opportunities, organized by the Panel on Confidential Data Access for Research Purposes, Committee on National Statistics (CNSTAT), Washington DC, 16-17 October 2003, revised draft for comment 1 December 2003, pages 36-39, <http://www.uwm.edu/%7Emargo/govstat/WS-MAcnstat.pdf>, quoted with permission.

xv William Seltzer and Margo Anderson, “Government Statistics and Individual Safety: Revisiting the Historical Record of Disclosure, Harm, and Risk”, prepared for a presentation at a workshop, Access to Research Data: Assessing Risks and Opportunities, organized by the Panel on Confidential Data Access for Research Purposes, Committee on National Statistics (CNSTAT), Washington DC, 16-17 October 2003, revised draft for comment 1 December 2003, page 36, <http://www.uwm.edu/%7Emargo/govstat/WS-MAcnstat.pdf>, quoted with permission.

xvi William Seltzer and Margo Anderson, “Government Statistics and Individual Safety: Revisiting the Historical Record of Disclosure, Harm, and Risk”, prepared for a presentation at a workshop, Access to Research Data: Assessing Risks and Opportunities, organized by the Panel on Confidential Data Access for Research Purposes, Committee on National Statistics (CNSTAT), Washington DC, 16-17 October 2003, revised draft for comment 1 December 2003, page 23-32 and 34-36, <http://www.uwm.edu/%7Emargo/govstat/WS-MAcnstat.pdf>, quoted with permission.. Also see Ellen Percy Kraly, “Blankets, Brass Tags and Bungalows: The Role of Population Data Systems in Historical Aboriginal Affairs in Northern Territory, Australia”, paper to be presented at the International Union for the Scientific Study of Population, XXV International Population Conference at Tours, France on 18-23 July 2005, <http://iussp2005.princeton.edu/download.aspx?submissionid=50449>.

xvii Australian Bureau of Statistics, *Information Paper: House of Representatives Standing Committee on Legal and Constitutional Affairs Inquiry into the Treatment of Census Forms – Submission from the Australian Bureau of Statistics, 18 July 1997*, <http://www.abs.gov.au/websitedbs/D3110122.NSF/0/a956800a923efbdf4a25667300095d05?OpenDocument>.

xviii Australian Law Reform Commission’s *Privacy and the Census*, Report No. 12, 1979, pages x and xv, also paras 1, 5 and 14, <http://www.austlii.edu.au/au/other/alrc/publications/reports/12/>. Also see Commonwealth Parliament, House of Representatives, Standing Committee on Legal and Constitutional Affairs, “A history of the current practice”, Chapter 2 in *Saving our Census and Preserving our History*, report, tabled 25 May 1998, <http://www.aph.gov.au/house/committee/laca/Inquiryincensus.htm>.

xix Commonwealth Treasury, *Destruction of Census Forms*, Press release no. 28, 1 June 1971, discussed in Australian Law Reform Commission’s *Privacy and the Census*, Report No. 12, 1979, at para 66, <http://www.austlii.edu.au/au/other/alrc/publications/reports/12/>. Also referred to in Commonwealth Parliament, House of Representatives, Standing Committee on Legal and Constitutional Affairs, *Saving our Census and Preserving our History*, report, tabled 25 May 1998, para 2.14, <http://www.aph.gov.au/house/committee/laca/Inquiryincensus.htm>.

^{xx} Australian Bureau of Statistics, *Information Paper: House of Representatives Standing Committee on Legal and Constitutional Affairs Inquiry into the Treatment of Census Forms – Submission from the Australian Bureau of Statistics, 18 July 1997*, <http://www.abs.gov.au/websitedbs/D3110122.NSF/0/a956800a923efbdf4a25667300095d05?OpenDocument>, where the ABS states: “Recent research commissioned by the ABS, and conducted by AGB McNair after the 1996 Census, indicates that there is significant public opposition to the retention of census forms. Key findings of the research are:

- 89% of respondents agreed that "Census forms should be destroyed to protect people's privacy and confidentiality";
- 67% disagreed that "Census forms should be stored for release in future for research purposes";
- Between 34% and 45% said they would be less likely to complete a Census form if forms were kept for release at some time in the future;
- Between 38% and 49% said the information on the Census form would be less accurate if forms were kept for release;
- 73% disagreed that "Researchers should be given access to Census forms including names and addresses".

“These findings are consistent with the qualitative evidence obtained from ABS consultations with the community about the census, feedback from census collectors and other empirical indicators of community attitudes to privacy and confidentiality, such as the increasing proportion of people and households opting to use privacy envelopes.”

^{xxi} Lynette Clemetson, “Homeland Security Given Data on Arab-Americans” (30 July 2004) *The New York Times* (abstract available at <http://query.nytimes.com/gst/abstract.html?res=F20815FA3A590C738FDDAE0894DC404482>); and see the defence by US Census Bureau Deputy Director, discussed in Hermann Habermann (Deputy Director, US Census Bureau), *Ethics, Confidentiality, and Data Dissemination*, 2005, <http://unstats.un.org/unsd/ethics.pdf>.

^{xxii} Kenneth Prewitt, “Additional Comments by Census Bureau Director Kenneth Prewitt on the Role of the Census Bureau and the Use of Census Tabulations in the Internment of Japanese Americans in World War II,” email to U.S. Bureau of Census staff, 24 March 2000, quoted in US Census Bureau, *Census Confidentiality and Privacy: 1790-2002*, 2002, page 16 <http://www.census.gov/prod/2003pubs/conmono2.pdf>.

^{xxiii} Hermann Habermann (Deputy Director, US Census Bureau), *Ethics, Confidentiality, and Data Dissemination*, 2005, page 1, <http://unstats.un.org/unsd/ethics.pdf> (footnotes omitted).