

**Submission to the
Australian Bureau of Statistics
on its Discussion Paper
*Enhancing the Population Census:
Developing a Longitudinal View***

10 June 2005

Robin Banks
Chief Executive Officer

Patricia Randal
Principal Policy Officer

Anne Mainsbridge
Solicitor

1. Introduction

The Australian Bureau of Statistics ('ABS') has invited public submissions in relation to its Discussion Paper, *Enhancing the Population Census: Developing a Longitudinal View* (Discussion Paper 2060.0, 26 April 2005).

The Discussion Paper indicates that the ABS is proposing to enhance the value of data from the 2006 Census of Population and Housing by combining it with future censuses. The central feature of the proposal is the creation of the Statistical Longitudinal Census Dataset ('SLCD').

This submission sets out the response of the Public Interest Advocacy Centre ('PIAC') to the changes proposed by the Discussion Paper. It is to be read in conjunction with the submission form completed electronically that was completed by PIAC on 9 June 2005.

It should be noted that, at this stage, this is PIAC's preliminary submission only. PIAC will make a further submission once the Privacy Impact Assessment Report—currently being prepared by a privacy expert for the ABS (Discussion Paper, p ix)—is made publicly available. In the absence of that report, PIAC is unable to comment comprehensively on the potential impacts of the proposal on privacy and confidentiality. As a 'fundamental aspect of this proposal is the management of privacy' (Discussion Paper, p 17), PIAC will comment in detail on the privacy aspects of the proposal after the release of the report of the Privacy Impact Assessment, taking the opportunity to comment also on that assessment. PIAC understands that the report will be available in late June 2005.

1.1 About the Public Interest Advocacy Centre

PIAC is an independent, non-profit legal and policy centre located in Sydney. PIAC provides legal advice and representation, public policy programs and advocacy training to promote the rights of disadvantaged and marginalised people and to enhance accountability, fairness and transparency in government decision-making.

PIAC specialises in undertaking matters that have systemic impact. Its clients and constituencies are primarily those with least access to economic, social and legal resources and opportunities. PIAC provides its services for free or at minimal cost.

Wherever possible, PIAC works co-operatively with other public interest groups, community and consumer organisations, Community Legal Centres, private law firms, professional associations, academics, experts, industry and unions to achieve its goals. PIAC works on public interest issues at a NSW, national and international level.

PIAC was established in July 1982 as an initiative of the (then) Law Foundation of New South Wales, with the support of the NSW Legal Aid Commission. Financial support for PIAC comes primarily from the NSW Public Purpose Fund and the Commonwealth State Community Legal Centre Funding Program. PIAC generates approximately forty per cent of its income from project and case grants, seminars, consultancy fees, donations and recovery of costs in legal actions.

1.2 PIAC's interest in privacy matters and expertise in privacy law

PIAC has a long-term interest in, and concern about, privacy issues within both the public and private sectors. PIAC's work as a legal and consumer advocacy organisation, particularly in relation to health matters, has required it to consider privacy issues as these frequently arise as matters of concern to many people who seek the Centre's assistance or involvement.

PIAC has provided advice and representation for clients in matters concerning alleged breaches of the *Privacy and Personal Information Protection Act 1998* (NSW) ('**the PPIP Act**') and the *Privacy Act 1988* (Cth) ('**the Privacy Act**'). Recently, PIAC represented the applicant in the matter of *MT v Director General, NSW Department of Education and Training*. This matter, which was heard by the Appeals Panel of the Administrative Decisions Tribunal on 31 May 2005, concerns the interpretation of several key provisions of the PPIP Act. The decision of the Tribunal has been reserved.

PIAC has played a lead role in health privacy debates in Australia in recent years, contributing to numerous inquiries and reviews at the national level. PIAC made submissions to the Senate Inquiry into Privacy and Private Sector in 1998 and to the NSW Health Information Privacy Review in 2001. In 2004, PIAC made a joint submission with the Blue Mountains Community Legal Centre to the NSW Attorney General concerning the review of the PPIP Act.

PIAC has provided much-needed co-ordination for consumer representatives involved in health privacy issues and development of electronic health records in NSW. The Centre has also been actively involved in various electronic health record initiatives of the Federal Government, including the national medication records developed as part of the Better Medication Management Scheme and the *Privacy (Private Sector) Amendment Bill 2000*.

2. General Comments

PIAC does not support the proposed changes to the Census outlined in the Discussion Paper.

PIAC's general concerns in relation to the proposed changes are set out below.

2.1 Consultation Process

The proposals outlined in the Discussion Paper envisage sweeping and dramatic changes to the nature of the Census and the manner in which information is collected. PIAC is concerned that the consultation process in relation to the proposed changes has allowed insufficient opportunity for detailed and careful consideration of changes of this magnitude.

In particular, PIAC is concerned about the short amount of time that has been provided for community consultation about the proposed changes. The Discussion Paper was released on 26 April 2005 and users of census data and other interested parties were given until 10 June 2005 to make submissions. Although the ABS conducted seminars in capital cities for interested parties and members of the public, information about these seminars was not made available until 3 May 2005 and insufficient notice was provided of the dates of these seminars. In late May, PIAC contacted the ABS to express interest in attending a seminar, only to be advised that seminars in Sydney had already been completed.

PIAC is also concerned about the lack of availability, at this late stage, of a Privacy Impact Assessment report. This makes it difficult for interested members of the public to have a clear understanding of the potential impacts of the proposals.

PIAC also notes that it was originally intended that the Australian Statistician would make an ‘in-principle decision’ on whether to proceed with the proposal or any part of the proposal in late June 2005 (Discussion Paper, p 27). It would appear that this timetable is no longer feasible or appropriate as the Privacy Impact Assessment report has not been made available prior to the deadline for submissions and is not expected to be available until the end of June 2005.

PIAC understands that interested parties, including the ABS, will be provided with further opportunity to comment on the report after it is released, and that there is also to be further time for advice to be sought by the ABS from the Australian Statistics Advisory Council (‘ASAC’) on the privacy aspects of the proposal. As the Australian Statistician will then need to fully consider information the submissions made by interested individuals and organisations as well as the recommendations from the Privacy Impact Assessment and ASAC, it is difficult to see how a decision about whether to proceed with the proposals could reasonably be made before the end of the year. Assuming that the Australian Statistician does decide to proceed with the proposal, it appears likely that there would then be a need for legislative amendment in order to allow for implementation of the proposals.

In PIAC’s view, there is not sufficient time to allow for detailed consideration and effective implementation of the proposals prior to the next Census, which is to be conducted in August 2006. PIAC believes that any process that attempts to implement the changes prior to that date is likely to be perceived by the Australian public as ill-considered and rushed, and is likely to be met with resistance and confusion. This may lead to severe loss of public confidence in the ABS, which may in turn result to a lower participation rate and less comprehensive and reliable data.

Recommendation

PIAC recommends that the ABS defer the impact of any further decision on the proposals to the 2011 Census and ensure the community is provided with clear and comprehensive information on the proposal, and an opportunity to then make submissions in the intervening period.

2.2 Use of data by ABS and other organisations without permission

PIAC understands that the ABS does not intend to seek consent from those participating in the census to use their data in the SLCD. PIAC further understands that this is because self-selection is seen as having the potential to destroy the statistical validity of the data as it is anticipated that many people would self-select out and, as a result, there would not be a comprehensive data set.

During the period of census processing, name and address information would be used to bring together census data and other selected datasets. This data includes:

- ABS Household Surveys;
- the 2005-06 Agricultural Census;
- birth and death register data (including cause of death);
- long-term immigration data; and
- national disease registers.

This would create a significant set of data about individuals that could be added to over time without any consent from individuals. Although the data would be de-identified, it would be matched using characteristics such as date of birth, sex, geographic region and country of birth. The address data would be grouped into lots of 200 households.

This data could then be used in three ways:

- quality studies, to help the ABS understand and evaluate the quality of its statistical operations and outputs;
- analysis of the 2005-2006 Agricultural Census data with Population Census data; and
- specific statistical analysis, where there is insufficient information to bring the data together with the SLCD using statistical techniques, to produce a dataset of adequate quality for a significant statistical purpose.

It is intended that such data will be de-identified and made available for research—on a user-pays basis—to researchers from both government and non-government, including private for profit, organisations. Any data on an individual who the ABS considers could be identified by unusual characteristics would not be supplied.

PIAC is concerned about the kind of organisations that would use this data and how they would use it, especially as it will have been collected without individual consent to use and disclosure.

According to the Discussion Paper, data will only be supplied for projects that have gone through an approval process. However, there are no details supplied about the criteria for the approval process, and commercial use of the data is not ruled out.

It is possible that banks, insurance companies and other commercial organisations could use de-identified personal information data for commercial gain through targeted advertising in particular geographical areas of 200 households.

Such detailed data is commercially attractive. It has recently been revealed that pharmaceutical companies have been buying de-identified data from doctors to be used for research and marketing (Frenkel 'Doctors selling medical records' *Herald-Sun*, 25 May 2005).

A further concern is that many of these organisations already hold name, address, date of birth and other data for many individuals. There is the possibility that they could cross-match this data with ABS data sets of 200 households to build enhanced data sets on identified individuals.

One of the examples given in the Discussion Paper of a government agency that may seek access to the data is the Department of Immigration Multicultural and Indigenous Affairs ('**DIMIA**'). According to the Discussion Paper, DIMIA is interested in tracking what happens to new arrivals in various visa categories (Discussion Paper, p 11). Again, there are no guarantees about how this data might be used in practice. Recent revelations about DIMIA's processes do not inspire confidence.

The use and disclosure of such information without individual consent appears to be in breach of certain of the National Privacy Principles. In addition, it is contrary to the expectations and wishes of most Australians. According to the March 2004 Roy Morgan research into the privacy attitudes of Australians, commissioned by the Office of the Federal Commissioner of the Privacy Commissioner, 64% of respondents felt that an individual's permission should be sought before de-identified information, derived from personal information about them, is used for research purposes.

2.3 Privacy Concerns

PIAC is concerned that there will be inadequate protection of individual privacy under the proposed arrangements. The proposed changes mean that there will be a substantial increase in the amount of data held by the ABS about each individual. Until now, there have been mechanisms in the census process that protect privacy. For example, there is a guarantee of early de-identification and there has been no association of individual records that arise from the Census with any other personal

data. The proposed changes appear to weaken some of these protections and deliberately cut across others. There is therefore greater potential for serious breaches of privacy to occur.

There is a disturbing lack of detail in the Discussion Paper about any proposed enhancement of privacy protections for what would be a vast increase in the amount of data that would be held by the ABS about individuals. It is to be hoped that the Privacy Impact Assessment will provide a comprehensive analysis of the privacy impacts of the proposal as well as proposals to ensure protection of the data.

PIAC will provide further, detailed comment on the privacy aspects of the proposals after consideration of the Privacy Impact Assessment report.

PIAC notes, however, that a major concern relates to the likely *perception* by the Australian public that their privacy will not be protected under the proposed arrangements. In PIAC's view, unless the Australian public can have complete confidence that the proposed arrangements will not threaten their privacy, they are likely to be reluctant to respond accurately to the Census. This is likely to impact on data integrity and quality (see 2.4 below).

2.4 Potential Impacts on ABS Reputation and Integrity of Data

The ABS currently enjoys a high level of community support and trust. This is due to its reputation for placing high priority on managing the privacy of information entrusted to it. PIAC notes that the ABS is acknowledged as denoting world best practice in this area.

PIAC is concerned that the proposed changes to the Census and, in particular, the public perception of the impact upon individual privacy of the proposed changes, may result in many Australians being less willing to provide accurate information to the ABS. This in turn is likely to impact on the overall quality and integrity of ABS data. It would cause lasting damage to the vitally important public resource that is the ABS census data.

In addition, the proposed changes to the Census appear to indicate a change in the role and function of the ABS. The public is likely to perceive that the role and function of the ABS is changing from that of a data collector to that of a data controller. There is potential for the expanded data collection to be perceived to be a springboard to the creation of a national population database that will enable all Australians to be continuously monitored.

PIAC recognises the importance of ABS data for social research, and that any process that impacts negatively on the collection of this data necessarily undermines the quality of the data. This will ultimately flow on to undermine the quality of research findings that rely directly or indirectly on the data.

In addition, removal of identifiable data—as is proposed by the ABS in the Discussion Paper—may compromise the usefulness of the data. Data is more likely to be identifiable where it relates to easily recognisable minorities. This may result in data relevant to homogenous groups being available with the consequence that minorities will not be appropriately reflected in research.

2.5 Lack of Certainty about Key Aspects of the Proposed Changes

PIAC is concerned that key aspects of the proposed changes are not sufficiently clear or certain. We note, for example, that it is unclear who will get access to the de-identified data. As outlined above, there appears to be nothing to prevent this data being used for commercial purposes. PIAC is concerned that there are currently no guidelines in place (or proposed) as to who would get access to this data and how they can use it. There is therefore a risk that it could be used for commercial purposes, resulting in people becoming the recipients of unwanted, targeted advertising. This is of

particular concern in view of the fact that this data will have been collected without individual permission.

Vague, unspecified terms are used throughout the Discussion Paper. These include:

... during the period of **census processing** (Discussion Paper, p 13)

Use ... likely to be of **community benefit** (Discussion Paper, p 21)

... **statistically appropriate** (Discussion Paper, p.23)

PIAC's specific concerns about these terms are addressed elsewhere in this submission. However, PIAC is also concerned that these apparently key concepts are not adequately defined in the Discussion Paper. In PIAC's view, these terms should be more clearly defined before the proposal can be seriously considered.

We also note that the process for evaluation, review and monitoring of the Census changes is as yet unclear and uncertain.

3. Comments re Specific Proposals

3.1 Proposal 1: Creating the SLCD

According to the Discussion Paper (Discussion Paper, p 4) the ABS is proposing to create the Statistical Longitudinal Census Dataset by:

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

The proposed method of bringing the census data together over time—referred to in the Discussion Paper as 'probabilistic record linkage' (Discussion Paper, p 6) would involve bringing together data from two different databases using a number of characteristics, such as age/date of birth, gender, geographic region and country of birth. All possible linkages based on these data items would be evaluated and the records for which the linkage is most likely to be correct would be brought together (Discussion Paper, p 6).

PIAC does not support Proposal 1 so far as it relates to 2006 data being combined with data from future censuses. The reasons for this are as follows:

- Individuals will become part of the SLCD without their consent to such usage.
- It is possible that individuals may be able to be identified without name through the probabilistic record linkage process. The Discussion Paper specifically acknowledges that:

For many cases, the two records brought together using probabilistic record linkage would reflect data for the same individual. (Discussion Paper, p 24)

There is therefore a very strong chance of the data being correctly associated with an individual, both in the form in which it is used by the ABS and in the form in which it is later provided to other users. The effect of this is that the data is not effectively de-identified.

- Individual concern about the possibility of being identified may cause individuals to subvert the Census process. For example, by not responding to it or by providing inaccurate information.

- Where there is an incorrect linkage, for example, because there is insufficient or incorrect information, this is likely to impact on the integrity of the data and undermine the quality of any research based on the data.
- Individuals will be given no opportunity to correct incorrect linkages or errors in matching.

PIAC does not support Proposal 1 so far as it relates to 2001 Census data being combined with data from the 2006 Census. The reasons for this are as follows:

- It is a potential breach of existing privacy laws and a breach of trust by the ABS as the 2001 data was provided in the 2001 Census on the understanding that it would be used for statistical purposes only. It was not the expectation or understanding of anyone who responded to the 2001 Census, that any data that they provided would later be used for other purposes and that it might subsequently be able to be related back to them.
- Concerns about the privacy implications of this proposal may undermine the reputation of the ABS.
- Concerns about the privacy implications of this proposal may lead to people subverting the Census process by not responding to it, or by providing inaccurate data. This is likely to undermine the quality of ABS Census data over time, and to ultimately affect the quality of research based on that data.

3.2 Proposal 2: Using the SLCD with other ABS Datasets

The ABS is proposing to make the SLCD available for use with other ABS statistical collections, such as ABS household surveys, using statistical techniques. It is proposed that use will be for statistical purposes only, with no identifiable data released outside the ABS (Discussion Paper, p viii).

PIAC does not support Proposal 2. The reasons for this are as follows:

- The Discussion Paper only provides detailed commentary about the use the SLCD with ABS Household Surveys (Discussion Paper, pp 8-9). However, it is clear from other sections of the Discussion Paper that it is also envisaged that ‘other collections conducted by the ABS’ could be involved (Discussion Paper, p viii). No information is provided about what these ‘other collections’ are. There are therefore no clear limits on the data-sources that may potentially be tapped into in order to expand the information contained in the SLCD.
- Use of the SLCD with these other datasets is to be for ‘specific approved projects only’ (Discussion Paper, p viii). However, no clear guidelines are provided as to the approval process. PIAC notes that one of the criteria for approval is that use has to be ‘likely to be of community benefit’ (Discussion Paper, p 21). This is a very flexible test and could be used to encompass use for law enforcement and social control.

3.3 Proposal 3: Using the SLCD with other Specified Datasets

The ABS is proposing to make the SLCD available for use 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. According to the Discussion Paper, use would be for statistical purposes only, and no identifiable data would be released from the ABS.

PIAC does not support Proposal 3. The reasons for this are as follows:

- Under the current proposal, it is envisaged that the SLCD will be made available for use with a limited number of datasets. However, it is extremely likely and there is nothing preventing the possibility that the number of possible datasets will increase over time.
- Use of the SLCD with these other datasets is to be for ‘specific approved projects only’ (Discussion Paper, p viii). However, as noted above no clear guidelines are provided as to the approval process.

3.4 Proposal 4: Using name and address information at the time of census processing

During the period of census processing, name and address information could be used to bring together census data and other selected datasets (Discussion Paper, pp 13-16).

PIAC does not support this proposal for the following reasons:

- This represents a new, unprecedented use of personal data by the ABS. Previously name and address information was only used for administration of census collection activities.
- It is likely that name and address data will be kept for longer than was previously the case, therefore increasing the potential for serious inadvertent or deliberate breaches of privacy to occur. The ‘period of census processing’ is a very flexible concept and is now likely to last longer than in previous censuses, because under the proposed changes processing would now include merging of census data with data from other datasets.
- The fact that name and address data will eventually be deleted is of limited comfort, as it will already have been used to link data with other sets of data.
- This proposal appears to be inconsistent with ABS’s responsibility to gather data only for statistical purposes.
- The ‘selected datasets’ include the birth and death registries, national disease registers and long-term immigration data. Some of these datasets arise from compulsory reporting to state and territory governments. Individuals who have provided this information have therefore not consented to the use of this information in the manner envisaged by ABS.
- The datasets are likely to multiply over time, for example, it is very likely that marriage registers will be added in order to track changes of name. There is no process in place to prevent this or to allow for consultation about the datasets to be involved.

3.5 Privacy and Confidentiality

PIAC is not in a position to make detailed comment about these issues until the Privacy Impact Assessment report is released. However, PIAC is concerned that certain aspects of the proposals appear to be in breach of existing privacy laws, and certainly at odds with the spirit and intention of those laws.

PIAC looks forward to providing comment upon the potential privacy impact of the proposals after the release of the Privacy Impact Assessment report.

PIAC recommends that the ABS convene meetings with interested organisations and individuals after the release of the report in order to allow for detailed discussion and consultation regarding the key aspects of the report. PIAC seeks to be involved in such consultations.

4. Conclusions

PIAC acknowledges the public interest in the collection and analysis of statistical information and the importance of having a comprehensive and accurate bank of Census data available for research. However, this needs to be achieved in ways that are consistent with promoting the public interest in protection of the privacy of personal information.

While it would appear that the creation of an SCLD could result in a rich source of information that may be used to inform research, in PIAC's view, it is vital to weigh the potential benefits of the SCLD against any negative impacts that its creation and use may have on the privacy of Australians.

PIAC is also concerned that any perception of infringement of privacy associated with the proposed changes may undermine the well-deserved high status of the ABS and affect the integrity of its data. The deficient consultation process associated with the proposed changes, coupled with the lack of clarity about key features of the proposal, do little to allay public fears about these issues.

Finally, PIAC's comments are, at this stage, of a preliminary nature only. PIAC looks forward to the opportunity to make further comments following the release of the Privacy Impact Assessment report. PIAC would also welcome the opportunity to be part of any discussion or consultation process concerning the findings of that Assessment.