



HOUSE OF REPRESENTATIVES

STANDING COMMITTEE ON LEGAL AND CONSTITUTIONAL AFFAIRS

Reference: Treatment of census forms

CANBERRA

Thursday, 23 October 1997

OFFICIAL HANSARD REPORT

CANBERRA

HOUSE OF REPRESENTATIVES
STANDING COMMITTEE ON LEGAL AND CONSTITUTIONAL AFFAIRS

Members

Mr Andrews (Chair)

Mr Barresi	Mr Mutch
Mrs Elizabeth Grace	Mr Randall
Mr Hatton	Mr Sinclair
Mr Kerr	Dr Southcott
Mr McClelland	Mr Tony Smith
Mr Melham	Mr Kelvin Thomson
	Mrs Vale

Matter referred to the committee for inquiry into and report on:

The treatment of forms from future population censuses.

The inquiry will examine the issues surrounding the destruction or retention of census forms. It will focus on whether the current practice of destroying census forms after processing should continue or whether they should be retained for medical, social and genealogical research.

WITNESSES

FOSTER, Mr Christopher, Assistant Secretary, Strategic Analysis and Evaluation, Department of Social Security, Athllon Drive, Greenway, Canberra, Australian Capital Territory	471
LEEPER, Mr Geoffrey, First Assistant Secretary, Department of Social Security, Athllon Drive, Greenway, Canberra, Australian Capital Territory	471
NICHOLLS, Professor Desmond Francis, President, Statistical Society of Australia, Department of Statistics and Econometrics, Australian National University, Canberra, Australian Capital Territory 0200	448
ROWLANDS, Mr David, Acting Assistant Secretary, Information Statistics and IT, Department of Social Security, Athllon Drive, Greenway, Canberra, Australian Capital Territory	471
SLESS, Professor David, Research Director, Communication Research Institute of Australia, 1st Floor, Old School Hall, Maitland Street, Hackett, Canberra, Australian Capital Territory 2602	465
TRUJILLO, Mr Gerardo Napoleon, Area Manager, Genealogical Society of Utah, 756 Pennant Hills Road, Carlingford, New South Wales 2118	457

HOUSE OF REPRESENTATIVES STANDING COMMITTEE ON LEGAL AND
CONSTITUTIONAL AFFAIRS

Treatment of census forms

CANBERRA

Thursday, 23 October 1997

Present

Mr Andrews (Chair)

Mr Barresi

Mr Randall

Mr Hatton

Mr Kelvin Thomson

Mr McClelland

Mrs Vale

Mr Mutch

The committee met at 10.13 a.m.

Mr Andrews took the chair.

NICHOLLS, Professor Desmond Francis, President, Statistical Society of Australia, Department of Statistics and Econometrics, Australian National University, Canberra, Australian Capital Territory 0200

CHAIR—I declare open this public hearing of the committee of inquiry into the treatment of census forms and welcome witnesses, members of the public and others who are attending this meeting of the committee. The subject of this inquiry is whether the current practice of destroying name identified forms after the data is collected from them should continue. The committee has taken evidence to date in all state capitals. This is the fourth hearing in Canberra and we look forward to hearing from witnesses today.

I welcome Professor Nicholls from the Statistical Society of Australia. In what capacity do you appear before the committee today?

Prof. Nicholls—I am Dean of the Faculty of Economics and Commerce at the Australian National University, Professor of Statistics and Econometrics, and current President of the Statistical Society of Australia. I am here today in the capacity as President of the Statistical Society of Australia.

CHAIR—Although the committee does not require you to give evidence under oath, I should advise you that the hearings are legal proceedings of the parliament and warrant the same respect as proceedings of the House itself. The giving of false or misleading evidence is a serious matter, and may be regarded as a contempt of parliament. At the outset, I apologise for the delay, but we were held up by a division in the House. We are in receipt of the submission from the society of 22 July. Can I invite you to make some brief opening remarks.

Prof. Nicholls—The submission was put in by my predecessor. I took over as president of the society in late July. From the point of view of the society, we believe that the maintenance of quality of data is of the utmost importance to Australia as a country. Having been sent the submissions, I have spent the last few days going through them and it has been rather interesting to me. I do not think I have ever seen a set of submissions that are so polarised into two groups: firstly, the government agencies, which are obviously looking to maintain the status quo through the maintenance of quality of the data, which they regard as extremely important—certainly any professionals are going to push that line—and the others, researchers, medicos, genealogists, et cetera, who are looking to have name identified data for their types of research. It is a situation where you have two competing groups. What one has to decide, obviously, is what is best for the country.

The Statistical Society is quite adamant it does not want to compromise the quality of census data. It is just so important, as many of the submissions pointed out, with respect to electoral distributions, financial considerations, allocation of funds to state and territory level, local government and so on. Other people confirm that there is a fair bit of

anecdotal evidence that in those countries where they maintain names and addresses they do have problems with non-response rates. My personal experience as a consultant to private enterprise from the university over a period of 15 years is that when you are running surveys the non-response rate increases. As soon as people can be identified with their responses, there is an element of caution and non-response rates and the quality of the data deteriorates. There is evidence of that, and anyone who works in that area has come across those problems. There are sensitivity type problems with the data.

In a number of the submissions I went through there was reference—certainly the Australian Bureau of Statistics, and I think it was Social Security and maybe one other—to an ACNielsen-McNair survey. I took it on, being in the role I am, to follow that up. I rang McNair's to have a talk with them about their survey, and they informed me that they had already been here and there had been some questions about the way they had run their survey. I think they referred to push polling, where I would have been referring to leading questions rather than push polling, and they ran a second survey and faxed a copy of that down to me yesterday with the issues raised there.

After the last census they ran a survey and, as is right through the evidence, found that 89 per cent of those surveyed believed that census forms should be destroyed—there is this privacy issue and the maintenance of quality of data—and the most recent survey they did rearranging and changing questions gave that figure of 70 per cent. There is an argument put in their report that the fall off would have been that the first time it was done was soon after the census was fresh in everyone's mind and the census is now well down the track so you would expect some leakage, which is what has happened. So, based on those types of results, they had actually gone out to try to find out from the public—I assume it was the Bureau of Statistics commissioned it; I do not know, someone obviously did the first one—to try to get some actual data. I have not seen any evidence in the submissions—I read them all thoroughly—from the medical researchers, et cetera, as to whether they have got any comparative data to counteract those claims. I certainly could not see any. But one also has to consider, of course, those other sources of data from registrars of births, deaths and marriages and medical records. As we move into the IT age there is no doubt more and more records are being kept and they are going to become more and more available as time goes on in the medical areas and other areas. I am certainly happy to answer any questions or attempt to.

CHAIR—In the submissions there is reference to data quality. How is the data quality of the census measured?

Prof. Nicholls—Data quality in the census—usually through non-response rates; people do not respond to particular questions. I am just trying to remember. I have had a look at so many submissions. I think Social Security referred to the fact that when it came to questions referring to family and personal income they had the highest non-response rate, and that tells you people are very cautious. They will not answer those types of questions, and straight away you have to say then, 'Well, we have to consider problems

with the representativeness of the sample.' In the census it is good because it is just so large. But if one starts getting very large proportions of people not responding then your data quality is seriously distorted.

CHAIR—If, for example, there is a reduction in data quality in the census, why can that not be corrected using statistical techniques and normal adjustment methods?

Prof. Nicholls—Well, you need to know. You still have to have information to tell you which way to make the adjustments. You have to get that data. Somehow or other you have to chase up the data or try to track the data to find which way to adjust the data: do you adjust it up; do you adjust it down, and by how much? You still have to have statistical information to make those adjustments. You may be able to say, 'This looks wrong when we compare it to the last census. Because the last census was more confidential we believe that is more correct.' Then we will make an adjustment that is needed according to what we found from the last census.

CHAIR—Can you not do that with some sampling methods?

Prof. Nicholls—You can try to do it with the sampling methods, but it can be very difficult because to do that you still have to go back and ask those questions that are causing the problems, and once more your response rates or the information given may not be appropriate. Obviously if there are identifiers on the census—for example, if there were names and addresses there—you can say, 'Okay, I can whip those through. We can do a cross-check, run the tapes against the tax office tapes and check out whether they are telling us the correct information.'

If you are prepared to run a database or run a survey asking personal information and people were prepared to put names and addresses on, then you could certainly do a cross-check, although I do not believe the tax office would give you the information—maybe they will; I do not know what the secrecy provisions are—but you could certainly run it through with the tax office. Similarly in the mid-1980s, I was involved in a review of fraud against the Commonwealth where we were trying to cross-match for Social Security and DEET at the time Austudy and unemployment data. It was a matter of running computer programs or databases across to check to see who was or who was not pulling both. These are the sorts of procedures you would have to follow to try to get the degree of accuracy of your data.

CHAIR—Is there not an argument for saying that some of the data collected by the census now is most likely to be inaccurate—for example, the income questions? I cannot recall what the actual figure was, but I recall from reading recently the initial returns of the 1996 census that the level of income it revealed across Australia seemed to be a gross underestimation by any other measure of income across Australia. So the question, I suppose, is: there must be varying degrees of inaccuracy, if I can put it that way, depending on the questions asked in the census now. Is that going to be any worse?

Prof. Nicholls—Or? There you are making the assumption that the other measures are the correct ones.

CHAIR—That is right, but compared to, say, estimates of income provided by the tax office, where they have policing powers, if I could put it that way, then my judgment might be wrong, but my judgment would be to rely more on what the tax office said on the levels of income than on others.

Prof. Nicholls—I would certainly agree with that. But I guess there is an argument one would put: if there were names and addresses on those census forms that people knew were going to stay there then that income data would be even worse, and it would get worse, there is no doubt about that.

Mr RANDALL—What evidence do you have to say that?

Prof. Nicholls—It is just personal experience I have had with running surveys, mostly through private companies out there. Whenever we ask for information about income levels the non-response rate is high, which quite often makes personal income data virtually useless.

CHAIR—At what level of inaccuracy does the data that is collected become so much less useful that it might create a problem? We have been provided with, for example, the underremuneration rates in censuses of population and housing in Australia, New Zealand, Canada, the US and the United Kingdom which range from a net undercount of 1.6 per cent in Australia, 1.2 per cent in New Zealand, 2.9 per cent in Canada, 1.6 per cent in the US and 2.2 per cent in the UK. Obviously there is a reasonable range. Is even the worst count—the 2.9 per cent in Canada or the 2.2 per cent in the UK—so unreliable that one cannot make reasonable judgments that one would wish to make for a whole variety of planning and administrative purposes, even on those where the undercount is the greatest?

Prof. Nicholls—I cannot give a definitive answer on that. It depends on what you want to use the data for and how acceptable estimates given to you are. I have heard anecdotal evidence, for example in the UK—this could be checked out, I am sure—that the 1991 census over there was so bad that they had trouble using the information from that to gain an estimate of the population. They were looking to go back to the 1981 census and project it forward. The order of accuracy that is acceptable is always dependent on what you want to use that data for. For example, if you are looking for an electoral redistribution, then how accurate do you want your data?

CHAIR—I have to chair the Main Committee for a little while. I will hand over to the Acting Chairman.

ACTING CHAIR (Mr McClelland)—Let us take as an option the possibility that

for the year 2001 census—because it is at the turn of the century—an experiment was conducted whereby the public were notified that it was the intention to retain census forms to become available the following century, but including a provision on that census form whereby someone could opt out of having their form retained. Would that do anything to alleviate these concerns that you have?

Prof. Nicholls—Provided people were aware of what they were doing on the day of the census. If you have as the first question, ‘Are you prepared to have your name and address recorded with this information for posterity?’

ACTING CHAIR—You are suggesting an opt-in rather than an opt-out option?

Prof. Nicholls—Once more, you come to your leading questions. Do you regard it as a leading question? If you put a question on the census, ‘Are you happy with the fact that names and addresses are destroyed now?’ or ‘Are you happy to have them left on?’ It depends how the question is worded. But the evidence from these McNair surveys is that more than two-thirds of the population are going to reject the idea of maintaining or keeping the data with the name and address identified.

ACTING CHAIR—The McNair survey, though, was a sampling. A test such as I have hypothesised there would be an actuality where you would test the general population.

Prof. Nicholls—Yes, but McNair are a professional organisation. They would have chosen their sample to be representative of the population. The sample has been chosen to be representative across age groups, states, et cetera.

ACTING CHAIR—In your view would some sort of opting out clause alleviate any of the concerns—‘Yes’, ‘No’ or ‘Maybe’?

Prof. Nicholls—I think there would have to be a massive marketing campaign beforehand to make the public aware that this was what was going on.

ACTING CHAIR—And to notify them that they have that option if they so desire?

Prof. Nicholls—That is right. You just could not dish up a census form and let them tick or cross the appropriate question at the time. I think there would have to be a significant marketing campaign to make the public aware of what it is about.

Mr BARRESI—In your experience, and through exposure to other research, would any particular socioeconomic group be more susceptible to having their data compromised by having their names on it? Would one group be more prone to it than another—lower income or lower educated groups? Would you be able to make that sort of assumption?

Prof. Nicholls—I do not think you can make that sort of assumption. I am just thinking quickly on my feet: at the one stage you have anecdotal evidence from talking to collectors of people refusing to put their names and addresses up-front now because they do not want their wives, ex-wives or de factos chasing them. On the other end, you have people in high tax brackets thinking, ‘All this information can be cross-checked with the ATO. We are going to have the Big Brother syndrome there.’ So there is nothing to stop them. As soon as you start asking about personal income levels and they have a name and address identified there, at some stage there is going to be that perception—the perception is the problem—that the information could be used to cross-check with, for example, the tax office; for unemployed people, the claims they make can be cross-checked with Social Security. There is just this keeping of tags.

Mr BARRESI—It seems to me that your major concern is about questions on the census relating to income. More than any other question, those questions create a lot of anxiety?

Prof. Nicholls—It is not just those, it is the quality of the data overall. I am sorry that I did not have a look at the most recent census form. I would have to go through and have a look at all the questions on housing, house ownership and personal assets. But the one that has always struck me in the marketplace is that as soon as you have anything to do with income that is where people back off. That is true of small business and business too, I might add.

Mr BARRESI—The ABS assured us the last time that there are a number of sources that are used in the planning decisions that are made—the census is just one. In your mind, would there be a value in having a census that does not have that highly sensitive question of income included because we can make those assessments through other sources anyway?

Prof. Nicholls—Every question would have been debated by a committee when they were forming the census questions. There would be very good reasons for having them there. I am sure that they would never put in questions where they could not get the information elsewhere. I do not know. Every question would have been scrutinised closely by an expert committee. If they were there, they would have had to have been there for a reason. I understand what you are saying, but maybe the income is a cross-check—it is a census; it is not a sample—of the accuracy of other sampling methods, for example. You do have your state-by-state comparisons and regional comparisons. I have not been involved, but my understanding is that a lot of committee work goes on in the setting up of the questions for the census. They would have to be there for a reason if they could not get the data elsewhere.

Mr RANDALL—I am going to follow on a little bit more from where Mr McClelland was coming regarding the 2001 survey. I believe from the evidence given to us—I have not been to all the inquiries but I have been to a large number of them—that

the marketing of this census as a point in history would be crucial to its success. For example, even the latest McNair survey was still operating under a shroud of paranoia built up by the Australia card—that sort of mentality. You get a different response. With a good advertising campaign many people would know what was happening, the reasons why, and that a point in time survey may be historically in the national interest. The point I am heading towards is that you are saying that you are concerned about the accuracy: there is only one way to find out and that is to hold a survey in 2001 and check how accurate the census was. If you have any problems, can you not project from the previous surveys and maybe one after to know the variance or the degree of error? Would that not be a good way to check how you were going statistically, in any case?

Prof. Nicholls—You could certainly have a look at it as a point in time in a set of longitudinal data. But how are you going to know the 2001 survey is inaccurate? With respect to the population, the information is going to be used for the distribution of grants and electoral redistribution databases. If it is out violently and there is a massive electoral redistribution and five years later you find out that it was absolutely wrong because of the last survey, you have a real problem.

Mr RANDALL—On that point, and I am certainly not a mathematician, but you are a statistician, I cannot understand why if you can project things accurately in other forms—you are saying, ‘Oh, but we are worried about this particular one where we could not be accurate’—you are saying that you cannot in this case?

Prof. Nicholls—The degree of accuracy of the projection is the problem. If you say, ‘I can be 95 per cent confident that the population is going to lie between 17 million and 17.2 million,’ that is great. But if all you can say as a result of bad data is that you are 95 per cent confident that the population is going to lie between 16 million and 19.6 million, what use is that range for you? It is when you get less accuracy in there that your estimates become more volatile. So you have to be more cautious with your estimates.

Mr RANDALL—But you are not dealing with pure science now. Maths is a pure science. This sort of survey is not pure. For example, I know that my electorate will change between elections—in three years—at least by 25 per cent. The electoral roll is out of whack to that extent between elections. We are dealing with people who do not bother putting in their changes of address or who falsify them—all these sorts of things.

Prof. Nicholls—Yes.

Mr RANDALL—For example, at the last election I had 71,000 people in my electorate. Now I am going to have 83,000, 81,000 or thereabouts. Somebody else in Western Australia had 93,000. I do not know whether that was so bad, other than the fact that the ones with the large number of people had to expend more resources. You are trying to say that, because you do not know to the last few people, this is a terrible set of information. I am saying—I am sure that I could be wrong—that there is a reasonable

degree of margin for error in these sorts of things because you are not dealing with a pure science.

Prof. Nicholls—But, for example, if you have two marginal electorates in outer Brisbane and the margins are pretty crucial for the redistribution, you want to be as accurate as possible there.

Mr RANDALL—This seems to be the axe that the bureau is hanging over us.

Prof. Nicholls—No.

Mr RANDALL—Can I say, with respect, that Mr McLennon from the bureau, who was here the last time, seemed to be pushing this argument that, if do you not behave yourselves, you might get redistributed out.

Prof. Nicholls—I wish I had been a spectator.

Mr RANDALL—Yes, you should have been. It is fun. I am just saying that this is the line.

Prof. Nicholls—He was a good rugby player in his day, let me tell you.

Mr RANDALL—I played Australian Rules.

Mr BARRESI—You look more like a rugby player.

Mr RANDALL—The point is that you are running this line. I am sure that it is your own, but we have had others do this. I find it a bit gross that this is the line from the bureau in particular; that it is trying to stand over us in this committee with that sort of threat.

Prof. Nicholls—I am sorry if it has come across like that. I do apologise. The line I am running is that, from the point of view of the Statistical Society and any professional statistician, you have to do your utmost to get the best quality data available. Data quality is crucial.

Mr RANDALL—My last question then is: you do not think that the Canadian information is quality?

Prof. Nicholls—My understanding is that the Canadians have now gone to running their censuses by removing names and addresses. They have been so concerned about it, as has New Zealand, that they have now removed the names and addresses to get that confidentiality. Internationally, there is no doubt that the Australian data is held in very high regard. I am sure others have given evidence—Mr McLennon and his colleagues—

along that line. That is certainly recognised internationally.

Mr MUTCH—How do Canada, New Zealand and the United Kingdom determine their electoral boundaries?

Prof. Nicholls—I am sorry, I could not answer that. I just do not know and I do not know how often.

Mr MUTCH—To date they have retained their name identified census records. I believe that Canada has a very high reputation with respect to statistical collections and dispute information. I believe that the reason they decided not to retain the last census—and this does not indicate future policy—was on the basis of storage. I do not think they have actually come to a resolution about future retention. New Zealand is also a bit interesting in that respect as well. They made a decision on the last one and they have periodically and intermittently kept them.

Prof. Nicholls—I am surprised they are doing that because of storage. In this electronic age, why could it not all be scanned in and kept electronically? I do not understand the logic there.

Mr MUTCH—With respect to the survey, your view on the reliability of the data is basically based upon ABS opinion?

Professor Nicholls—No, that is not true. We have had a survey methodology group within the Statistical Society. We have a conference every two years. We always have a one-day workshop on survey and survey methodology, usually looking at current research, theoretical concepts, et cetera. There is some practical work there, but the whole idea of that is refining techniques to get the best quality data possible. That is what the society is about.

Mr MUTCH—But you have not conducted your own surveys on the question of whether, if we kept name identified records, a reliable census could be taken?

Prof. Nicholls—No, certainly not. But I am sure that, if someone was prepared to pay, we could organise it.

ACTING CHAIR—Professor Nicholls, thank you very much for giving us your time this morning.

[10.46 a.m.]

TRUJILLO, Mr Gerardo Napoleon, Area Manager, Genealogical Society of Utah, 756 Pennant Hills Road, Carlingford, New South Wales 2118

ACTING CHAIR—In what capacity do you appear before the committee today?

Mr Trujillo—I am representing the Genealogical Society of Utah, which is the arm of the Family History Department of the Church of Jesus Christ of the Latter Day Saints, better known as the Mormons.

ACTING CHAIR—Although the committee does not require you to give evidence under oath, I should advise you that hearings are legal proceedings of the parliament and warrant the same respect as proceedings of the House itself. The giving of false or misleading evidence is a serious matter and may be regarded as a contempt of parliament. That is a statement we make to all witnesses. We have received your submission. Would you care to make some introductory comments?

Mr Trujillo—I would be happy to. Thank you very much for allowing me to come to this beautiful place and to present our case regarding the necessity for us to retain the census data records. I am not quite sure whether I need 45 minutes to make my presentation and to make a positive argument as to why the Australian government should make every effort to save census data, not only the most recent one in 1996, if it is not too late, but also future census data. I can only hope that I can convey this message today to the committee as to how strongly we feel about this subject. I am speaking not only on behalf of myself or on behalf of the Genealogical Society of Utah but also for over 100,000 patrons who use our facilities throughout Australia only. If we were to include the patrons who use our facilities worldwide, it would come to many, many millions.

As I said to you in the beginning, my name is Napoleon Trujillo and I was born in Quito, Ecuador. For those honourable members of the committee who do not know where Ecuador is, it is in South America. The equator line passes through Ecuador, and I imagine this is how it took the name Ecuador. About 27 years ago, I came to live in this country and I have not regretted it ever since. I arrived on a beautiful sunny day in September 1971. I live in Castle Hill in Sydney, and I have a wife and five beautiful children.

In April 1973, I started to work for the Genealogical Society of Utah. The first thing that I want to do is give you my credentials and what we do to establish the reasons why this petition came to you. The Genealogical Society of Utah was established about 102 years ago in Salt Lake City, Utah, with the specific goal or objective of assisting the members of the Church of Jesus Christ of the Latter Day Saints to find their genealogy, their roots and ancestors.

Also, in 1934, a large microfilming project began, and now the Genealogical Society of Utah is the possessor of over two million rolls of microfilm. Sometimes I am a little bit hesitant to use the name of the Genealogical Society of Utah because, in the past, I have received correspondence with some interesting addresses. I have one letter that I had framed in my office which reads 'To Mr Napoleon Trujillo from the Gynaecology Society of Utah'. I thought that was interesting. So from now on we refer to ourselves as the Family History Department. Some people, when you refer to genealogy, seem to think that it refers to some very highly technical field. So the Family History Department is now the trade name that we use. But because the Genealogical Society of Utah has been around for 102 years, a lot of the record holders of this world know us better as the Genealogical Society of Utah, so we still use that.

I started microfilming records in Australia at the Mitchell Library in Sydney. Then I moved to the Dixon Library in the same building. One of the first projects that I had the opportunity to do was the filming of the 1828 census. I did not know the importance of the census records then. From there I moved to the state archives of New South Wales, where a substantial project took place. We microfilmed the convict records of Australia. Many thousands of volumes were microfilmed. For some of you who would know, these are available now through many outlets throughout Australia. These records have been preserved and are being used by researchers on a daily basis.

After three years in Sydney, I moved to Hobart, Tasmania, and started to microfilm the records of the state archives of Tasmania. We also had a project to microfilm the convict records, the land records, the probate records and any other records that came to our knowledge. While I was there I was also invited by the Uniting Church of Tasmania to microfilm their baptisms, marriages and burials. We stayed in Tasmania for 12 months. Shortly after that, I had some brief stints in microfilming a number of records all over Australia. I went to microfilm the probate records in the Supreme Court in Perth. I have also microfilmed the Supreme Court records in Queensland and the Supreme Court records in Townsville and Rockhampton. Shortly after that, I moved to Adelaide and microfilmed the records of the state archives. I went to the local libraries there and also some of the churches. I have filmed many records in New South Wales, most notably the two that I mentioned prior, and, of course, many local shire councils, Newcastle Library and so on and so forth.

During the past 25 years we have also microfilmed a large collection of births, deaths and marriages from the different church denominations. We have microfilmed the diocese of Grafton, Wagga Wagga, Armidale and so on. In 1987, I was appointed to be the regional manager for the South Pacific operations, with additional responsibility to oversee the projects in the Pacific islands, which included Australia and New Zealand.

In 1995 I was appointed to be the area manager for this part of the world, which included Asia, with added responsibilities to oversee the microfilm projects in the Philippines, Indonesia, Sri Lanka and India. Some of the largest projects that we have ever

undertaken have taken place in those countries. To give you an example, we have recently completed a film of the civil registration of the births, deaths and marriages of the Philippines. That took us almost 10 years, employing about 250 people. We now have in our possession about one hundred million frames of the vital statistics of the Philippines. A similar project has been undertaken in Sri Lanka and India, but to a lesser extent. The same thing could be said of Indonesia.

At the present time, we are embarking on a very large project in the Philippines to microfilm the Catholic records of the country and the census records that are kept in Manila.

ACTING CHAIR—Mr Trujillo, I appreciate you giving your credentials. It is very interesting to hear what can be done and it is important for us to know that the technology is available. However, we are on a tight time schedule and we are running very close to time. It may be better to dive straight into questions and then you can tidy up anything at the end of it.

Mr Trujillo—Sure. I was trying to explain that I am uniquely qualified to tell you what researchers are looking for, especially those who use our family history centres in Australia and worldwide. The project that I mentioned is but a small drop in the collection that we have done.

ACTING CHAIR—If Australian census records were retained on, for instance, microfiche and were kept confidential for 100 years and were then released to the public, would they contain information which would not otherwise be available to researchers through any other source?

Mr Trujillo—No. There would be duplication under births, deaths and marriages and other records, but they would have a unique value to the genealogist. Every little bit of information, such as the street and suburb that somebody lived in and who else lived in the house, is of utmost importance to genealogists and family historians. I was going to get to that point.

Let me give you one example of where we place the census in our collection in terms of its importance. Recently, we undertook a most massive project to index the 1881 census of England. I do not know whether anybody has mentioned this to the committee. We found the data contained in the census of 1881 to be most important. Until I read the report that I have here, I did not understand the massive nature of the project.

Let me give you a couple of details. It took nine years to complete. In excess of 30 million records were indexed. It took 1,500,000 hours or about six years just to complete the initial transition from paper to computer. About 9,000 individuals were involved in the work of transcription, auditing and the production of the end product. About 170 computers were purchased and they worked day and night to complete the project. Many

challenges and obstacles had to be overcome. My head spins just thinking of the gigantic and humungous task that was undertaken. Despite all that, the 1881 census that has been released to the public—first in microfiche format and eventually it will be available in CD-ROM format—has proved to be one of the most wonderful tools that genealogists, family historians, sociologists, demographers and medical teams can use to research data and formulate a number of research papers.

We place a great deal of importance on the census. There is no question about that. It is sad to think that, as you look at our collection of data, in the Great Britain area you will find that we have microfilmed or indexed the 1676, 1901, 1841, 1851, 1861, '81 and '91 censuses. When you look at our collection, sadly, only one census has been kept. I could tell you categorically, based on our experience as to the usage of the records, that out of the 70,000 rolls of microfilm that are used within Australia, between 20 per cent to 25 per cent of requests are requests for census data.

ACTING CHAIR—And roughly, how many would that be?

Mr Trujillo—Twenty-five per cent of 70,000.

Mr BARRESI—I think we have pretty well monitored why geologists place a lot of faith on the information that is provided through some of the censuses, particularly the 1891 census. You have mentioned already that part of your job is to microfilm lots of other data that is available in the community. You talked about Catholic records, Bathurst, Townsville and Wagga Wagga. That information is already there. In 1891 other source information was very limited, apart from perhaps christenings or baptismal records that churches held. There are other avenues for genealogists to collect information. Forget 1891; in modern times, why is the census an important component of your work, considering the other information that is available?

Mr Trujillo—The first point is the amount of information that it contains uniquely in this format. Some of the information that is found in the census will be found nowhere else. When I talk about making it available now, I believe that, if the decision is made to keep the records, we would want them to be released in about 70 to 100 years time when privacy issues will no longer be as important as they may be today. For genealogists, every bit of information, every relationship and every paper that is found is of the utmost importance to link families.

Mrs VALE—You have explained some of the uses that your census records have been put to by the various groups. Would you like to elaborate on that any more?

Mr Trujillo—For instance, in the United States where each state has a census and various censuses are kept, many medical teams have researched the censuses to establish genetic disease linkage data. Demographers have used the different censuses for various reasons, not necessarily only related to the redistribution of boundaries of politicians.

Mainly, that has been for scholarly research.

Mrs VALE—Your organisation is funded by the Church of Jesus Christ of the Latter Day Saints?

Mr Trujillo—That is correct.

Mrs VALE—Do the users pay any fees to access your records?

Mr Trujillo—There are two systems. Firstly, the records are already here in Sydney. We have a large collection in a large warehouse where we have 100,000 rolls of microfilm. If those records are available and you come to Carlingford, there will be no charge. If we have to ship them to Perth, Darwin, Adelaide, Brisbane, there will be a charge for the postage.

Mr BARRESI—Why has the church given itself such a mission to collect these records? What is the purpose for the church in doing this on behalf of the wider community? I can understand it doing for its own members and followers, but why do it for the wider community?

Mr Trujillo—We asked the same question ourselves. About 80 per cent of the people who come through our facilities are not members of our faith. We believe that we have received a commandment to preserve the records of the world and to make them available to the people. This only became a reality about 25 years ago. Before, we had a single-minded mission to assist the members of the church.

We opened the family history centre in the church facility. We installed microfilm readers, computers and the like. Then, as we opened that to our members, all of a sudden genealogy blossomed tremendously and people started to come to our facilities. We cannot get them out of the door. We have had to accept the fact that people come. As genealogy has increased in popularity, people come to research their roots and this has created a boom in family history. Also, we sponsored the world preservation records in 1970 in Salt Lake City and out of that there was a worldwide boom in family history. So we see the function of the church to preserve the records of the world; we have microfilm in 100 countries of the world and we provide those facilities to anybody who comes to our facilities regardless of creed, religion, colour, whatever.

Mr BARRESI—You mentioned in your previous answer to me that the census is an invaluable source of information to which all genealogists look to get some sort of historical snapshot and to make certain decisions. The ABS and a few others have already told us that having named census data will compromise the integrity of the data. Surely data whose integrity has been compromised is of marginal value to a genealogist. Therefore, of the data you have available to you in some of the other locations where named data is available, what is your assessment of its use to those who are accessing that

information?

Mr Trujillo—I am not quite sure exactly to what extent the ABS has indicated that the data has been comprised.

Mr BARRESI—That is an assumption because obviously we have never had it before.

ACTING CHAIR—We do not currently preserve our census.

Mr Trujillo—Yes, I know.

ACTING CHAIR—The bureau has expressed concern that if we did it would be compromised.

Mr Trujillo—We have found that a lot of people have made some mistakes in church records and the vital statistics. A lot of people change their names, yet somehow the information is still available to the public. But, by finding the source, even if it was compromised, is a starting point to a family historian rather than having to find the correct data eventually.

Mrs VALE—One of the things that you pointed to was the preservation of records. That is one of the issues that the committee is going to have to look at—the storage of such records if it is decided that we should keep census forms. Could you explain how your microfilming process works and why have you chosen that as the format for your records?

Mr Trujillo—The technology of microfilm has been around for quite some time. It has been proven that if it has been kept at an ideal temperature and humidity control it will last for a very, very long time. There is an archival standard that has been chosen by many of the archives and libraries of the world. It is not the most technologically advanced piece of preservation; nevertheless, we have two million rolls of microfilm in our collection. We have a sizeable investment in it. We have in Salt Lake City in the granite mountain vault the ideal conditions to keep our collection. We could have decided to, for instance, scan the records but that technology has not been proven; it has not been able to be aged long enough to find out if we can take a laser data diskette in about 50 years to see if it has lost its value or if the data has been erased.

Mrs VALE—Out of curiosity, what kind of format were you using before microfilm?

Mr Trujillo—Nothing, it was just the gathering of volumes and records, then microfilming came. We believe that microfilm, even though it seems to be an outdated technology, is still very much an important tool for us.

Mrs VALE—What is the cost factor in microfilming? Do you have any idea?

Mr Trujillo—It depends. We have mobile teams whereby we do not move the records out of the localities. For instance, when we went to microfilm the diocese in New South Wales we had a team of camera operators who went to the actual churches and parishes and microfilmed the records in the premises themselves.

Mrs VALE—So you have never actually costed it?

Mr Trujillo—We have a lot of volunteer labour; we call them missionaries. If I was to give you a rough cost per microfilm, it would be about \$100 to produce one roll of microfilm.

Mr BARRESI—I have just a brief question. If we decided to retain the named census data, the Australian Archives has a capacity to store the information; we have already been out there. Would the church still be looking at microfilming the data and having its own copies; if so, why, considering it would already be there in the Australian Archives?

Mr Trujillo—No, we will not look at microfilming. If that format eventually comes into the public domain—either in the archives, through genealogical societies or whatever repositories—if it is decided in the future that it will be available, no, we will not. Why duplicate something that has already been done? But if the national archives was not going to do anything, then we will certainly come and make a proposal to them that we will at least microfilm the records.

Mr RANDALL—Just following up on that, are you aware that that information might not be available for 70 to 100 years?

Mr Trujillo—Yes.

Mr RANDALL—That does not change your answer then?

Mr Trujillo—No, because I am a record preservation man, I would like to see the records preserved rather than destroyed.

Mr RANDALL—I have a brief question which follows Mr Barresi's question. It has been suggested, as Mr Barresi said, by a few others that retaining name identified census data would compromise the quality. Do you consider yourself eminently qualified to answer that question and, if you do, what is your answer?

Mr Trujillo—No, I am not qualified to answer that categorically. I have lots of experience as to the usage of records by the many thousands of people who come. They are welcome to have the data even if it has been compromised. I hope that it has not. I do

not see any reason why it should.

ACTING CHAIR—The thrust of your evidence is that this data is very valuable, it is something which will be used and the facilities are available to preserve—the technology is available to perform the task. Is that a fair summary?

Mr Trujillo—Yes, I will add my voice to the many other people who have come before this committee to indicate to you that we should make every effort to keep the census data. It is of utmost importance. We value that data and any other data, very much so.

ACTING CHAIR—Thank you very much for your time. We appreciate you coming along.

[11.10 a.m.]

SLESS, Professor David, Research Director, Communication Research Institute of Australia, 1st Floor, Old School Hall, Maitland Street, Hackett, Canberra, Australian Capital Territory 2602

ACTING CHAIR—I welcome Professor Sless. In what capacity do you appear before us today?

Prof. Sless—I am the Research Director at the Communication Research Institute of Australia and I am here in my capacity as the research director.

ACTING CHAIR—I should state formally that, although the committee does not require you to give evidence under oath, I should advise you that the hearings are legal proceedings of the parliament and warrant the same respect as proceedings of the House itself and that the giving of false or misleading evidence is a serious matter and may be regarded as a contempt of parliament. I understand that we have not received a submission from you, but would you like to give us an overview or some introductory remarks as to what you would like us to consider?

Prof. Sless—I am here in a technical capacity rather than to address the substantive issue that the committee is trying to decide. I was asked to look at some of the material that is before the committee—surveys which purport to point to the attitudes of the Australian public on the keeping of the census forms. It is in that capacity, therefore, that I am before the committee. I have had a chance to look at the original survey that was done by AGB McNair and the more recent survey that was done by ACNielsen-McNair subsequent to that first one. I think the recent one is submission 276.

I should say at the outset that, within the limits of what is possible in professional market research, the ACNielsen-McNair survey is a proper and professionally conducted survey within those limits. However, I think it is important to understand that there are inherent weaknesses in that type of work. Therefore, using evidence from that type of work to arrive at decisions is a risky business. In summary, that would be my view of that data as it is before the committee.

If I can elaborate on that, generally speaking there are three tests one applies to any kind of survey, any kind of test material: that is, its validity, its reliability and its sensitivity. If we were to take at face value the issue of reliability, we could say that there is a reasonable degree of reliability between what the first survey measured and the second survey measured—reliability is whether, if you take a measure on one occasion and you repeat it on another occasion, you substantially collect the same data. That is what reliability is about. The question of validity goes to the heart of whether you have asked the right questions. The question of sensitivity is whether you have covered the subject matter with a degree of nuance that is appropriate to the subject itself.

Those are the three tests one would apply. In normal scientific survey methodology, there are procedures that ensure or try to ensure that those three things are done. In the case of commercial market research, these are sometimes used but not used entirely. In this case, I would have to say that in the first survey they were used inadequately and, in the second survey, they were not used at all, unless one wanted to regard the first survey as a pilot study in preparation for the second survey that was done.

It is a standard piece of survey methodology that one runs pilot studies. For example, if you look at the ABS census, they go through a very elaborate procedure in the development of their collections. They not only do pilot studies, they do observational studies of people completing the form and they also do what they call dress rehearsals. Some of these rehearsals involve many thousands of subjects. In other words, every effort is made by them to ensure the accuracy and validity of their data. In the case of commercial market research, very rarely is there the time or the type of budget that would allow this type of work to be done. Therefore, it simply is not done. I think that there is little expectation on the part of people who commission this type of research for that kind of level of sophistication to be applied. As I say, that is not a criticism of the company that has done the research; it is simply a statement of fact about the nature of this type of work.

Having said that, there are some deeper underlying problems and they have really to do with notions of attitude and attitude measurement. If I can briefly summarise these for you. Typically, the assumptions on which surveys are based that purport to measure attitude are that, first of all, attitudes are isolable, that is, you can actually isolate a thing called an attitude and you can measure it. Secondly, there is a close link between attitude and behaviour. So knowing what attitudes people have will tell you something about what people will actually do; in other words, it will enable you to predict their behaviour. Thirdly, and this is often assumed particularly in relation to things like public information campaigns, if you can change attitudes then you may also be able to change behaviour. In other words, if you bring people's attitude around to your way of thinking, then you may also change their behaviour. Of course, the other factor is that attitude surveys are, indeed, instruments for measuring attitude. Those are the kinds of assumptions on which this type of methodology is based.

As I say, that is the conventional wisdom of our time. As a result of some of my own research and a collaboration between myself and Ruth Shrensky of the University of Canberra, we have been looking at the background to attitude research and attitude theory and have come to the conclusion—and, by the way, this is material that is shortly to be published and available in the public domain—that, first of all, there is no general agreement on what attitude actually is. If you go to a researcher, you are likely to find conflicting views of what attitude actually is. There are no commonly agreed views of what attitude is. Secondly, the evidence shows that if you change attitudes, it rarely results in changes of behaviour. These things are not necessarily linked. Thirdly, there is no generally agreed method for measuring attitude. We do not know of any method that

reliably, sensitively or validly measures attitudes. Attitudes do not necessarily predict behaviour and communication activity, that is, running campaigns and so on does not necessarily result in predictable changes in attitude. Therefore, that suggests that for the type of evidence you have looked at in this case, whatever the opinions expressed there may be, I do not think that they provide a basis for judgment in this area.

ACTING CHAIR—Just on that last bit, even if the survey data does reflect an attitude, could the concerns of those who feared preservation of the data be alleviated if, for instance, on the next census if it was going to be preserved it had a question to enable the subject to opt out of having their records kept?

Prof. Sless—As I was saying earlier, the ABS has very elaborate methodologies for the development of the census material. I would see no reason why they could not actually find answers to these questions as part of that pilot investigation of the census itself. In other words, they do not need to go externally to get another study that would determine whether such a move to preserve the material would actually alter the quality of the data. They have the capacity and indeed the methodologies to find that out for themselves.

Mr MUTCH—If this government were to instruct the ABS to retain name identified census information in future, can you give us an indication of how this should be achieved or the modus operandi that should be followed?

Prof. Sless—It seems to me that the primary question that the ABS are concerned about is the quality and integrity of its data—the capacity to use that data for all the kinds of many planning purposes and other purposes to which it is put. I think that is their primary concern. If there was a real concern about privacy, then that should determine how they proceeded with the census. What I am saying is that I think they have the capacity in their development methodology to advise the government whether or not this is the case.

Mr BARRESI—I agree with the comment you made about attitude and predicability of attitude. However, some will say that an individual's past behavioural actions are a far better predictor of what that person will do in the future. Various interviewing techniques are based on that, such as behavioural questioning. Is it possible for an organisation, whether it be ACNielsen or whatever, to devise a questionnaire which rather than questions a person's attitude, questions that person's behavioural history?

Prof. Sless—People are notoriously unreliable in giving personal accounts of what they have and have not done in the past. This is not because they necessarily lie or forget. It is simply in the nature of the thing that, if you ask me to tell you what I did yesterday, I will construct a story about what I did yesterday. In other words, I will recreate the events.

ACTING CHAIR—Yes, but that is only valid if it is a one-off question which is not followed up with subsequent behavioural questioning.

Prof. Sless—Yes.

ACTING CHAIR—My comment would be: would it not be far more valid in terms of getting an assessment of people's views to engage in focus groups rather than a written questionnaire, where one is able to follow up with a series of structured behavioural questions?

Prof. Sless—The same applies to focus groups, in fact even more so, because you have group dynamic effects that can get in the way. But you are quite correct: the best predictor of my future behaviour is what I have done in the past. Therefore, your best way of determining what my future behaviour might be is to observe my behaviour, not my account of that behaviour. Our institute developed the guidelines for the collection forms that the ABS now uses. In that methodology, there are procedures for observing people completing forms. That enables you to predict what their likely behaviour will be in the future, based on observing their current behaviour. If that were done properly and sensitively by the ABS at the time it goes through its pilot studies for the census, it would be in a position to advise the government whether at that point the retention of names and addresses was a good idea—on the basis of evidence of behaviour, not on the basis of evidence from surveys.

Mr MUTCH—It has been said that those surveys indicate significant public opposition to the retention of census forms. Is that drawing too long a bow?

Prof. Sless—Yes, I think it is. I think even on the second survey there is an implied correct answer, if you like, to the critical question. If you have a look at the survey question—and I am thinking here of the first one; I understand that all of the issues are to do with the rotational order—you see that it states, 'Census forms should be destroyed to protect people's privacy and confidentiality.' I would be hard pressed to answer in the negative to that. It seems to me that that is not a question but an argument and a statement.

Mr RANDALL—Professor Nicholls said that rather than being push polling, the questions were leading. Would you agree with him?

Prof. Sless—Yes. That is normal practice.

Mr RANDALL—Is that why you are saying it would be harder to answer in the negative?

Prof. Sless—Yes.

Mr MUTCH—But the normal practice is to ask leading questions, is it not?

Prof. Sless—I think so. As far as I have seen from surveys, yes, it is very common; it is a commonly used technique.

Mr MUTCH—Is that because you are trying to get a result or because you are trying to do it properly?

Prof. Sless—I think people are trying to advise their clients appropriately, and I think they take the appropriate methodologies to do that.

Mr RANDALL—I wish to pursue that point. I have been involved in situations where consultants have been called in. For example, in local government, my cynical experience is that you get a consultant to give you the right answer so that you can present it to your council or whatever as a fait accompli. You keep going until you either get the right one or somebody that you know who will give you that answer. Given the almost intransigent position of the ABS on wanting to retain the census data, would you believe that this is what has led to this form of implied questioning?

Prof. Sless—Not necessarily. I think people have acted properly and with good faith and with a desire to find the evidence. I think the problem is, if you like, a structural one—that is, why ask the question in the first place? It is a bit like some of the research that was done in the US by Jensen on the relationship between race and intelligence. One has to ask at the beginning: why is that an interesting or important question to ask? In other words, there is an implied view in the question, that is, that race as a factor matters.

It is not so much that people do things in a biased way. The bias is already inherent in the question. That is very clear from that first question. That is not a question. It does not ask, ‘Do you think we should destroy the name and address on the census form?’ and then ask, ‘Why?’ It offers someone quite a strong argument and then says ‘agree or disagree’. It is not an open-ended question. In other words, from the range of acceptable social answers, you know in a sense when you are asked that question what the possible answers are. Protecting privacy is like motherhood. Why would you want to do something that did not protect privacy?

Mr RANDALL—In other words, it is a leading question?

Prof. Sless—Yes.

Mr MUTCH—We are often told that although name identified material has been kept in countries like Canada, New Zealand, the UK and so forth, Australia is so different from those countries that we could not do the same. Do you have any insights into that argument? Do you think that is a valid argument?

Prof. Sless—I think there is a certain ethical or social issue that, if you have proceeded by a particular set of rules, people have a right to expect you to continue to proceed with that set of rules. I do not think that is a research question that I could answer in any technical sense, but I think there is a certain moral issue at stake. Having said that, I repeat that the ABS does have the capacity to make the necessary investigations into this on its own behalf and in its own right. It has the methodologies and skills to do this. It would not add substantially to its own costs in the development of the census form to do so.

Mr MUTCH—As Mr Randall has said, we have had a lot of resistance to the concept from the ABS. In fact, I think it almost had to be dragged kicking and screaming to do that. I hope it would be professional enough to be able to do that. Is that the only body in Australia that has the capacity to do that sort of work?

Prof. Sless—The ABS has an impeccable reputation. Its sampling methodologies, its handling of sampling errors and so on, and many other aspects of its collection methodology are first class. I would be more inclined, as I think most researchers would be, to trust ABS data than almost any other social or economic data that is available. There is no question of that. I would not in any way suggest that they would bias it unnecessarily. It is possible, of course, to get them to make the methodologies that they use and so on, as they do in most of their surveys, a matter of public record and subject to peer review and proper scrutiny. So I do not see a problem with that at all.

Mr MUTCH—Is there anything further that you wanted to add?

Prof. Sless—No.

CHAIR—Professor, on behalf of the committee, thank you for coming along today and discussing these matters with us. A division has been called in the House, so I will adjourn the hearing until after the division.

Short adjournment

[11.47 a.m.]

FOSTER, Mr Christopher, Assistant Secretary, Strategic Analysis and Evaluation, Department of Social Security, Athllon Drive, Greenway, Canberra, Australian Capital Territory

LEEPER, Mr Geoffrey, First Assistant Secretary, Department of Social Security, Athllon Drive, Greenway, Canberra, Australian Capital Territory

ROWLANDS, Mr David, Acting Assistant Secretary, Information Statistics and IT, Department of Social Security, Athllon Drive, Greenway, Canberra, Australian Capital Territory

CHAIR—Gentlemen, although the committee does not require you to give evidence under oath, I should advise you that the hearings are legal proceedings of the parliament and warrant the same respect as proceedings of the House itself. The giving of false or misleading evidence is a serious matter and may be regarded as a contempt of parliament. We received the department's submission of 28 July of this year. Would you care to make some brief opening comments?

Mr Leeper—Thank you, Mr Chairman. I will be brief, because I think our submission is relatively straightforward. The Department of Social Security is a significant user of the services of the Australian Bureau of Statistics, including census material. Clearly, we have an interest in the quality and integrity of the data that ABS collects, including the census data itself.

Our submission focuses on the committee's first term of reference, which focuses on the quality and integrity of the data and the impact that retention of census forms might have on the quality and integrity of that data. Our submission states quite clearly and, having looked at the transcripts of the committee's hearings and other submissions, I should make this point quite clearly up front: we do not claim to have any strong evidence that retention of the census forms would lead to a diminution in the integrity of that data. Our submission is certainly couched in terms of saying, to the extent that retention of the forms undermines confidence in the integrity of the data holdings themselves and that that impacts on people's preparedness to respond to the questions at all or to some extent, then the department would have some concerns about a decision that would require the forms to be retained. We are not saying that we have strong evidence that retaining the forms would be a problem from the point of view of integrity, but to the extent that it is a problem and may undermine the data that the ABS collects and the integrity of that data, the department submits that that should be carefully considered.

CHAIR—Could you elaborate on what use the department makes of census data?

Mr Leeper—It has quite a wide range of uses. I will make a statement and then

my colleagues might help me out. We use it to provide information in areas where the department does not hold data. Clearly we have extensive data holdings of our own that are, I guess you would call, in the nature of a census because they relate to detailed information about individuals at a point in time. Of course, many Australians are not recipients of Social Security payments. In our advisings to government on social policy matters, we do make extensive use of census data in those areas where we do not have information concerning either individuals or classes of individuals in the Australian population.

Mr Foster—Examples of that would be in the areas of family policy where you might be looking at issues of take up of family payments. You need a base population, which is the sort of data you get from census or census-driven surveys to compare your own administrative data with. In the housing policy area, the same sort of issues arise. You need to put the context of your own Social Security clients into the broader national context. The census is often the only place where you can get detailed client/population and small area data which you need when looking at some of these issues. There are many uses in Social Security, because we cover a very broad range of policy interests. Social Security clients are an important part of the overall population, as are shifts that occur in that group. We can provide you with a written list of specific uses, but those are a couple of examples where it is clear.

CHAIR—In terms of small area data, in your submission you say that you require data on the characteristics of relatively small groups such as Aboriginal and Torres Strait Islander people or recently arrived immigrants. For your purposes, does the data that is collected on that come from ABS or do you rely more on what you collect yourself? Does the department itself do any surveys of small groups?

Mr Foster—The answer is all three. We do our own data collections in certain cases. We make extensive use of ABS data and we would rely on administrative data. No single source gives us the complete picture in the range of Social Security interests, so it just depends on the circumstances.

CHAIR—I take what from what you are saying, Mr Leeper, that you are not making a judgment nor have the evidence about whether or not there would be a diminution in the quality of the data if the census information was retained. However, have you made any judgment or assessment of what level of diminution would have an impact? You might answer that by saying that any diminution has some impact, but I am trying to find out when the diminution would have such an impact that it would make the sorts of estimates and judgments that the department has to make reach the point of unreliability?

Mr Foster—We have not done any assessment of that sort of thing.

Mr RANDALL—Before I make an observation, I ask the question: were you

asked to come here?

Mr Leeper—The decision to present a submission was a decision taken by the department.

Mr RANDALL—Were you contacted by the Australian Bureau of Statistics to make a submission?

Mr Leeper—It is certainly true that the Australian Bureau of Statistics wrote to our secretary on 3 June alerting us to the fact that the inquiry was being held. They included some information relating to the AGB McNair results. However, the decision to lodge the submission was the department's decision. The decision to appear here is in response to an invitation by the committee. As far as I am aware, there has been no contact between ourselves and ABS in relation to this matter, nor have we provided them with a copy of our submission to you.

Mr RANDALL—I think I have made my point. You have come here saying that you have not any evidence that there will be any corruption of the information, other than the information we have from the bureau. You really say, 'Look, we use it,' and it is very nice for us to hear that you use it. But there is really not a lot of reason for you being here other than to say, 'We use it and we would not like it to be corrupted.'

Mr Leeper—I think that is probably reason enough to warrant a submission and to respond to your invitation to appear. The census is important in two ways: one, for the data that it produces, which provides a very useful snapshot, point in time picture. It is also useful and is essential, as I understand it, to benchmark the other surveys that the ABS undertakes and of which the department is also a very heavy user—things like the household expenditure survey, the income distribution survey, the labour force survey. All of those things that are sample-survey based have their design, and in some respects their integrity, vested in the quality and integrity of the census data itself. It may come back to the point that the Chairman was making before: to what degree can people stop providing information relevant to their circumstances? If the forms are retained, at what point will that start to impact on the quality of the results? I cannot answer that. But, to the extent that the census database begins to either lack information or have incorrect information in it, the subsequent benchmarks that are derived for the purposes of other sample surveys themselves become incorrect and the sampling error just builds upon itself.

Mr RANDALL—But you have no hard-core evidence. I go back to your point about point in time surveys being important to the year 2001 census. Would it be a good point in time survey to retain, for some of the reasons that you have mentioned in terms of future historical data, international interest, and so on?

Mr Leeper—I do not have a strong view on that. Whether a decision is made to retain that material or not, our submission merely says that if that results in people being

concerned about the information that is being provided on the grounds that it will be kept, and that results in either not answering questions or giving information which is not full and accurate, then we have concerns about the quality of the data that might result from that census.

Mr RANDALL—Realistically—and I will finish on this—you are just repeating your old mantra over again, aren't you?

Mr Leeper—I am repeating the points made in the submission, certainly, yes.

Mr MUTCH—With respect to your note in your submission that census data is valuable to the department in advising government on social policy issues, do you have a research arm? Are you yourself, for instance, involved in research, or are any of the gentlemen before us involved in the research arm?

Mr Foster—My branch has an overall strategic responsibility for research and evaluation in the department. All areas of the department in each of the particular program areas with their particular responsibilities, whether it is families or the age pension area and so on, all have a direct use of the data.

Mr MUTCH—Then the better the research that you do, the better the advice you can give to the government of the day?

Mr Foster—Yes. The Department of Social Security is primarily a policy development department.

Mr MUTCH—We have heard evidence from sociologists from academic departments who suggest that if they had access to name identified material they could provide invaluable information on the basis of research to governments for specific longitudinal studies and other types of studies. So my question is: if the retention of name identified material had no adverse effect on the reliability of the statistics collected by the ABS, would that then be of benefit to your research arm?

Mr Foster—We do not have a view on that. I would think that the sorts of longitudinal research that the Department of Social Security would get involved in would probably be specifically funded longitudinal research. I am not sure that we have a current need to go back over long periods on individuals on the census data. It could be potentially useful, but we have not made our submission on that basis.

Mr MUTCH—Would you agree that that is on the basis that you have not had any access to any data of that kind anyway.

Mr Leeper—I think I would need to ask how long the exclusion period would be before the data became available, too.

Mr MUTCH—Some submissions have suggested that, for important academic research, the exclusion date should not apply at all, and for medical research and so forth.

CHAIR—If I understand the submissions correctly, that is on the basis that the researchers would not have access themselves to the name identified characteristics of the data but would, through a matching system, be able to obtain the data: that is, the privacy of the particular individuals would still be maintained, but the data would be available for longitudinal research. For example, I understand your department is responsible for the policy aspects—just to pick a couple off the top of my head, you mentioned family payments and, say, things like child support payments.

Mr Leeper—Yes, that is correct.

CHAIR—Presumably, at the moment the only way of tracking that in terms of research is to track those persons on any longitudinal basis who come onto your books, so to speak.

Mr Leeper—That is correct.

CHAIR—Presumably, though, if one wanted more complete research it would be useful to be able to track a group of, say, 100,000 people, as I understand can be done in the UK, and look at any time at what percentage of that group—let us take child support—were parents in a custodial or non-custodial role, how that changed over time, whether private arrangements are in place that you may not have details about necessarily from the information that you collect, and what proportion are doing that. It would seem to me—if I can put it this way—from a purist point of view, that that information would be extremely valuable in making longer term policy decisions but that some of that information at least is not able to be collected and analysed in Australia because we do not retain the census and use it in the way in which I understand it is used in the UK.

Mr Leeper—Yes. In relation to the case you mentioned of child support, because things are very dynamic in that area, there would be some value in a single point in time every five years, but it would also date extremely quickly. It would really be like a time slice. It would be very hard to pick up the dynamics of family formation and disformation, I guess.

CHAIR—But the way it is used in the UK, as I understand it, is that whenever this started 20 or so years ago there was a sample group. I stand to be corrected on the figure, but my recollection from previous evidence is that there was a group of, I think, 500,000 through name identification taken out as a sample group. That group obviously diminishes each year as a random sample because people die, emigrate or whatever and, therefore, it is topped up again, so you always have on a year-to-year basis an ongoing sample of half a million of the population. It is not just taking it as a snapshot every five years. That is what has been put to us. Mr Mutch was elaborating upon the use of name

identified data to be able to produce such a group which can then be maintained through topping up where it drops off, which would be extremely valuable for purposes of social research.

Mr Foster—I think that is a particularly interesting and useful survey. Social security would certainly have an interest in looking at ways in which you can have such longitudinal surveys. I guess the question we were responding to was whether we have a concern if retention of the census name identifiers would lead to a drop in data integrity. I think the survey you are talking about is a separate thing which people know they are going into and retain their records for that purpose.

CHAIR—It is taken from the census. That is where the pool came from in the first place; it was taken from the census. It is then provided in such a way so that the researchers obviously do not have the names; privacy is protected. But, unless you take it from the census in the first place, you do not get the pool of 500,000 or whatever the exact figure is.

Can I take this a step further? I am exploring this in my own mind. Again, using a hypothetical case in terms of child support, we know, for example, that the out-of-wedlock birthrate in Australia has increased substantially since the end of World War II from, I think, something like four or five in 100 to 26 or something like that in 100 on the latest figures. That can be broken down between those who were never married, where children are born out of wedlock, and I think the other major group obviously is those who have been married or in a relationship which, for whatever reason, fails and you have a group there.

How do you make projections into the future about, for example, child support payments? What do you base your projections in the future on in terms of behaviour of people entering into relationships, for example, where people are unmarried but a child results from that and the relationship does not proceed and, in the other group, where there is a relationship on foot—whether de facto or de jure—that does not continue? Given not only the amount of payment that is involved in this but the obvious emotional and other trauma and, at times, antagonism that we as members of parliament face in our electorate offices every week over this issue, how is that measured in terms of advice to the government about what is going to happen, say, over the next five years?

Mr Rowlands—I think one source of advice would be the ABS. Certainly the number of sole parents in the community and the trend in the numbers would be a valuable item of data in making those sorts of projections. They are derived from surveys done by the ABS relatively regularly. But those surveys themselves depend, as I understand it, on the census for their structure on the basis of sampling.

CHAIR—But that does not tell us on a month-to-month, almost year-to-year basis, anyway, what is happening in terms of behaviour, does it? Whereas if you had a group of

500,000 people who were regularly surveyed and you could see that a certain proportion were, for example, living together and then moving apart and a certain proportion were having a child, you would have much more accurate trends in the way things were going. As you said, what we do now is that we get a five-monthly census, we get the returns of births, marriages and deaths, which keeps that up to date, but it does not tell us much about what is happening in people's lives, their living arrangements and things like that which, from a policy point of view, may be important in making future projections.

Mr Leeper—I think we have from time to time done longitudinal surveys of our own customers. I acknowledge the point that you are making about the broader population. My memory of those studies is that there is quite a degree of volatility, particularly in the sole parent population, with relationship formation, break-ups and so on. The point that I think Mr Rowlands was making was that we use the ABS data to benchmark the starting points. We then monitor fairly carefully trends within our own populations and project those to try to provide advice to government. Those things can be measured over time. There will always be events that catch us by surprise in forecasting expenditure on different payments. We mainly rely on our own administrative data to provide the detailed forecasts to government about likely expenditure pressures or forecasts in that area. We consistently go back and check that our data makes sense against what is produced by the ABS in terms of broader population measures; otherwise from time to time you can get some silly situations where our data and the ABS data may not match for a range of reasons, definitional and otherwise.

Mrs VALE—You say in your submission that the question on income already has a substantial degree of sensitivity attached to it and that, as a result, it has one of the highest rates of non-response to all the questions asked. What is your view of the proposal that income-related and other sensitive questions be removed from the census to allay any possible public concern about the retention of the census?

Mr Leeper—I think it is a fairly important question from a measurement point of view and also with respect to the benchmarking work that the ABS does in other areas, particularly things like the income distribution survey and the household expenditure survey. From a departmental point of view, we would be concerned if that question disappeared because of its importance.

Mr Foster—A lot of social security work requires an understanding of what is happening in changes in income within the community. So to remove all of that would be of concern to us.

CHAIR—My understanding is that in terms of its accuracy, that question is notoriously unreliable. How much reliance can you place on it even now?

Mr Rowlands—I do not think that there are very many better sources. At least that gives us information on people on very low incomes who may never furnish a return to

the tax commissioner. We have data on our own customers, but when it comes to people who are not social security recipients—for example, if we are giving advice on policy that might include some new customers—that is the sort of data that we must use.

CHAIR—Do you check that in any way? For example, can you go to the census data for a census collection district or districts? Do you take that data at the lowest level and then run cross-checks against your own data that is provided to you through the various recipients of social security?

Mr Rowlands—We certainly use that low level, fairly detailed census district data for policy advice and things other than policy advice. It is probably worth pointing out to the committee that we can use this census data for monitoring implementation where, for example, we extend new payments to people who were not previously our customers. That sort of data can tell us where there are concentrations of potential recipients. We can target publicity, monitor how well the take-up is going and see just how effective the implementation has been.

Mr Leeper—I am not aware that the department's administrative regions are concordant with the ABS local government collection districts. That has been one of those longer-term things that everyone has wanted to get towards, that is, making sure that the administrative boundaries for different Commonwealth purposes are aligning themselves carefully with census collection districts. That would be a major advantage. My understanding is that that is not the case at present. So we cannot easily match our own administrative data down to the local level with ABS census output at a very, very small level. We can use it as a different view in a geographic area, but I do not believe that those things relate one-to-one with our own administrative boundaries.

CHAIR—If you were contemplating establishing a new office or relocating, would you use it for that purpose?

Mr Leeper—Absolutely. Demographic projection in that sort of area is very important to us and that is where we would use ABS information quite carefully, yes.

Mr HATTON—You have indicated that personal, family and household questions are very important. If it was the case that we looked at locking up that information for academic researchers but still making it available in the census data made available through the ABS, would that still suit your purposes?

Mr Leeper—By 'locking up', do you mean it would be carefully protected?

Mr HATTON—Yes. We could even go a bit further. Some of the critics have suggested that sensitivities over personal information are so great that, if you were going to allow researchers access to records, one of the ways to overcome that would be not just to have a matching exercise so that they had blind access to the information but to say

that particular questions on income, families, households and so on should also specifically be locked away from those academic researchers. But the corollary of that is that the public information available through C Data '96 or the version that will follow it would still be made available; it is just that the personally identifiable information would not be. How would that affect your situation?

Mr Leeper—I think that would depend on the extent to which individuals felt assured that the architecture was such that their information would be protected. There is a great deal of sensitivity attached to government holdings of information. Our own department is very aware of that and we take great pains to protect people's privacy as far as possible. My own view is that it would depend on the extent of assurances that were given to people that the data would be protected absolutely. It would take only one problem to undermine that. The only thing that I could think of there that might be relevant is that the further down you drilled into some pieces of information, perhaps even in remote areas, it may become possible to identify particular individuals or groups of individuals, even though you have locked away some of that personal data, just because there are not many people out there in some of the collection districts. That is an issue that would need careful attention.

Mr HATTON—That is possible now. However, the ABS scrambles the information at that CCD level. It combines that information from couples so that you cannot get to the person or people directly. It provides that type of scrambling.

Mr Leeper—The critical issue is confidence. If a decision is taken along the lines that you are talking about, the issue for me would be how we assure people, come census night in the year 2001, that the information they are being asked to provide will be protected in that way. That is the critical issue for me.

Mr HATTON—People in the financial sector have more and deeper information on every individual in the country. For example, when you apply for a bankcard, a MasterCard, a personal loan, or whatever, your whole financial history has to be laid open. Once that information is in the financial sector, there is no great guarantee, apart from general legislation, that it will not be traded around the country. In terms of what the public might perceive to be the situation, the privacy surrounding census information is probably far greater than it is with respect to information in the financial sector.

Mr Leeper—That is a fair point.

Mr HATTON—You have talked about household expenditure surveys and their importance to the department. All of that information is subject to how good the census is. How vital are they to you? Do they give you a very close picture of what the state of society is at any one point in time?

Mr Leeper—Let me try to put it in a policy context. There is a great deal of

speculation at the moment about possible changes to the mixture of the taxation system. In the event that the government decides to look at options for broader based consumption taxes, the information from the household expenditure survey will be absolutely essential, because it gives you consumption patterns that we can correlate with income levels to work out possible approaches to compensating people for the possible price increase effects of a broadly based consumption tax. It is that essential. Without that detailed information on spending by consumption area, by household and by income levels, we and other departments would struggle to advise the government properly on how to design appropriate compensation. That is indeed the policy intent of the government in relation to that particular matter. It would be very, very hard.

Mr HATTON—Over time, has it become much easier for your department to access the census information as it has been made available in more manipulative ways—for example, with C Data '91 and '96—or were you always in this position, having access to the mainframe information?

Mr Rowlands—We are certainly increasingly greater users of the C Data. The number of copies of it that we are obtaining from the ABS is increasing each time.

Mr HATTON—Can you see the information in better ways? I am thinking of MapInfo for Windows and what is done in Space of Time Research 1 as well. You can see the data graphically and also see the relationships between the data in a more intuitive way than by just running through the numbers. Has that type of thing been used?

Mr Rowlands—I believe so. We also tend to go back to the ABS and ask for specific tables if they are not available within the C Data material. That is often done. Alternatively, we specifically pay the ABS to do some further tabulations based on the census data for particular policy purposes.

Mr HATTON—But you are in a position to do that because you can see the situation better?

Mr Rowlands—Yes.

Mr HATTON—And you can ask better and more defined questions?

Mr Rowlands—Yes.

Mr Leeper—The technology has been a great 'enabler'.

CHAIR—Mr Leeper and Mr Rowlands, thank you very much for your department's submission and for discussing it with us today. I thank everyone who attended today and also those who recorded the evidence.

Resolved (on motion by Mr Hatton):

That this committee authorises publication of the evidence given before it at public hearing this day.

Committee adjourned at 12.19 p.m.