

1996 CENSUS DATA QUALITY: INCOME

Rosa Gibbs
Trevor Knight
Population Census Evaluation

December 2000

© Commonwealth of Australia 2000

This work is copyright. Apart from any use as permitted under the *Copyright Act 1968*, no part may be reproduced by any process without permission from AusInfo. Requests and inquiries concerning reproduction should be addressed to the Manager, Legislative Services, AusInfo, GPO Box 84, Canberra, ACT 2601.

In all cases the ABS must be acknowledged as the source when reproducing or quoting any part of an ABS publication or other product.

Produced by the Australian Bureau of Statistics

CONTENTS

1.	INTRODUCTION	1
2.	COMPARISON OF THE 1991 CENSUS INCOME QUESTION WITH THE 1996 CENSUS INCOME QUESTION	2
	2.1 Changes to the 1996 Question	2
	2.2 Comparison of Income Distributions, 1991 Census with the 1996 Census ...	2
	2.3 Non-Response Analysis	4
	2.4 Effect of Dummy Records on the Non-Response Rate	6
3.	NEGATIVE AND NIL INCOME, 1996 CENSUS	8
	3.1 Negative Income	8
	3.2 Characteristics of Persons Reporting Negative Income	9
	3.3 Characteristics of Persons Reporting Nil Income	10
4.	COMPARISON OF THE 1996 CENSUS WITH THE 1995-96 SURVEY OF INCOME AND HOUSING COSTS (SIHC), AND TAXATION STATISTICS 1995-96	11
	4.1 Comparison of Distributions of Income from the 1996 Census with SIHC ..	11
	4.2 Comparison of Distributions of Income from the 1996 Census with Taxation Statistics 1995-96	13
5.	THE 2001 CENSUS INCOME QUESTION	15
6.	CONCLUSION	16
	APPENDIX 1: 1991 Census Income Question	17
	APPENDIX 2: 1996 Census Income Question	18
	APPENDIX 3: 1991 and 1996 Censuses Income Ranges	19
	APPENDIX 4: Distribution of Income, Persons Aged 15 and Over, 1991 and 1996 Censuses	20
	APPENDIX 5: Taxation Statistics 1995-96, Taxable Individuals by Fine Grade of Taxable Income, Australia	21

LIST OF TABLES AND GRAPHS

Table 1:	Income Non-Response Rates by Sex, Australia 1986, 1991 and 1996 Censuses	4
Table 2:	Persons with Not Stated Income by Labour Force Status (LFS) and by Sex, Australia 1996 Census	5
Table 3:	Total Population Over 15 Years of Age by LFS, by Income Stated and by Income Not Stated, Australia 1996 Census	6
Table 4:	Effect of Dummy Records on Income Non-Response Rates, Australia 1996 Census	7
Table 5:	Characteristics of Persons Reporting Negative Income Values, Australia 1996 Census	9
Table 6:	Characteristics of Persons Reporting Nil Income Values, Australia 1996 Census	10
Table 7:	Distribution of Income for Persons Aged 15 and Over, 1995-96 Survey of Income and Housing Costs (SIHC) and 1996 Census	12
Graph 1:	1996 Census, Income Question	3
Graph 2:	1991 Census, Income Question	3
Graph 3:	Comparison of 1995-96 Survey of Income and Housing Cost (SIHC) with 1996 Census Income Data (Excluding Not Stated)	13
Graph 4:	1995-96 Taxation Statistics, Taxable Income	14

SUMMARY OF FINDINGS

The 1996 Income Paper evaluates the data quality of the income question in the 1996 Census. The issues analysed in this paper include: a comparison between income distributions in the 1991 and 1996 Censuses, non-response rates, the impact of the introduction of new low income response categories and comparisons of 1996 Census data with both the 1995-96 Survey of Income and Housing Costs (SIHC) and the 1995-96 Taxation Statistics.

The main findings of the paper are:

- The non-response rate decreased from 9.3% in the 1991 Census to 6.2% in the 1996 Census.
- The introduction of the new 'negative' and 'nil' income response categories had a positive impact on data quality from low income earners.
- The income question was not perceived by respondents as being as sensitive and intrusive, in particular by middle and high income earners, as had previously been thought. The response rate for 'middle' and 'high' income earners improved in the 1996 Census with 98% of employed respondents answering the question. The high response rate from this group suggests that 'middle' and 'high' income earners did not perceive the question to be intrusive.
- There is a general tendency by respondents to underestimate their income. In particular there is a higher non-response rate for persons not in the labour force, and a tendency by respondents not to include government and private transfer payments when stating their income.
- 1996 Census income data has proven to be comparable with a variety of other data sources, namely: 1991 Census income data, 1995-96 SIHC and 1995-96 Taxation Statistics. Despite methodological differences, the 1996 Census compared favourably with these series in its general trend at all income levels.
- The data will be useful for users as a comparative indicator of income levels across small geographical areas and small population groups.

1. INTRODUCTION

Income data from the census are important as a socio-economic measure, providing an indication of individual, family and household well-being between different population groups and in different areas of Australia. It is frequently cross-classified with other census variables to provide income 'profiles' of the population within small geographic areas and for small population groups. Income data are widely used by both government and non-government organisations such as Centrelink, the Department of Veterans' Affairs, welfare organisations, academics and other researchers. In particular, it is used to identify and locate the disadvantaged for the planning of support services.

A question on income was first asked in the 1933 Census in order to gauge the effects of the Depression. It was not included again until the 1976 Census but has been included in every subsequent census. As in previous censuses, the 1996 Census question on income required each person aged 15 years or more to indicate the range within which their gross income from all sources fell. Family income and household income were then derived from the individual income data. Testing has repeatedly demonstrated that asking respondents for income within a range provides better quality data than asking them to provide an actual dollar amount. This may be due to the difficulty of accurately calculating an exact dollar amount.

The main data quality problem associated with the collection of income data on the census form is the self-enumerated nature of the census (with no collector/interviewer assistance to help answer a complex question). Another problem is the space restrictions on the form that limit the number of income questions that can be asked. These issues lead to misunderstanding and misreporting by respondents, with the following impact on reported income data:

- relatively high non-response rate for the question;
- a perception by some groups in the population, particularly those respondents who are not in the labour force or are unemployed, that the income question does not apply to them as they do not earn an income. Consequently, such persons either do not respond to the question or select the 'negative' or 'nil' income response options; and
- the difficulty of estimating gross weekly income from all possible sources of income.

Despite these data quality concerns, there was a large decrease in the non-response rate for the income question in the 1996 Census. The proportion of non-respondents declined by a third, from 9.3% in the 1991 Census to 6.2% in the 1996 Census. Several contributing factors are discussed in this paper as possible explanations for the decrease in the non-response rate. There is also a comparison of income data distributions from the 1991 Census with the those from the 1996 Census.

2. COMPARISON OF THE 1991 CENSUS INCOME QUESTION WITH THE 1996 CENSUS INCOME QUESTION

2.1 *Changes to the 1996 Question*

Several changes were made to the 1996 Census income question. The major change was an increase in the number of response options from 14 categories in the 1991 Census to 16 categories in the 1996 Census. The two additional categories were 'nil' income and 'negative' income. These options were included in 1996 after the Australian Bureau of Statistics (ABS) Housing, Income and Expenditure Section recommended capturing positive, nil and negative income to bring the census into line with the newly developed ABS income classification standard (Census Working Paper 96/2: *1996 Census Form Design and Testing Program*). Although a 'nil income' category had been used in censuses prior to 1991 the use of a 'negative income' category was new.

The adoption of the ABS income classification standard also meant that an adjustment to the low income categories for the 1996 Census question was necessary. In the 1991 Census, the two lowest income response categories were less than \$58 per week and between \$58 and \$96 per week. For the 1996 Census the 'negative' and 'nil' income categories were added and the two lowest income response options were \$1 to \$39 per week and \$40 to \$79 per week.

The other change to the income question in 1996 was the reversal of the response categories from ascending to descending order. This was done so that the 'nil' and 'negative' income categories could be placed as the last two response options with a view to minimising over-response to these categories. Testing revealed that the inclusion of the 'nil' and 'negative' income categories and the reversal from ascending to descending order did not adversely affect the response distributions (Census Working Paper 96/2: *1996 Census Form Design and Testing Program*).

Other modifications to the 1996 income question included the repositioning of the question to the top of the page; in 1991 it was the second question on the page. The list of sources of income, in the question's instructions, was also slightly modified to include possible sources of income not included in the 1991 Census namely, 'rental assistance' and 'commissions and bonuses'. Copies of the questions used in both the 1991 and 1996 Censuses are included in Appendices 1 and 2.

2.2 *Comparison of Income Distributions, 1991 Census with the 1996 Census*

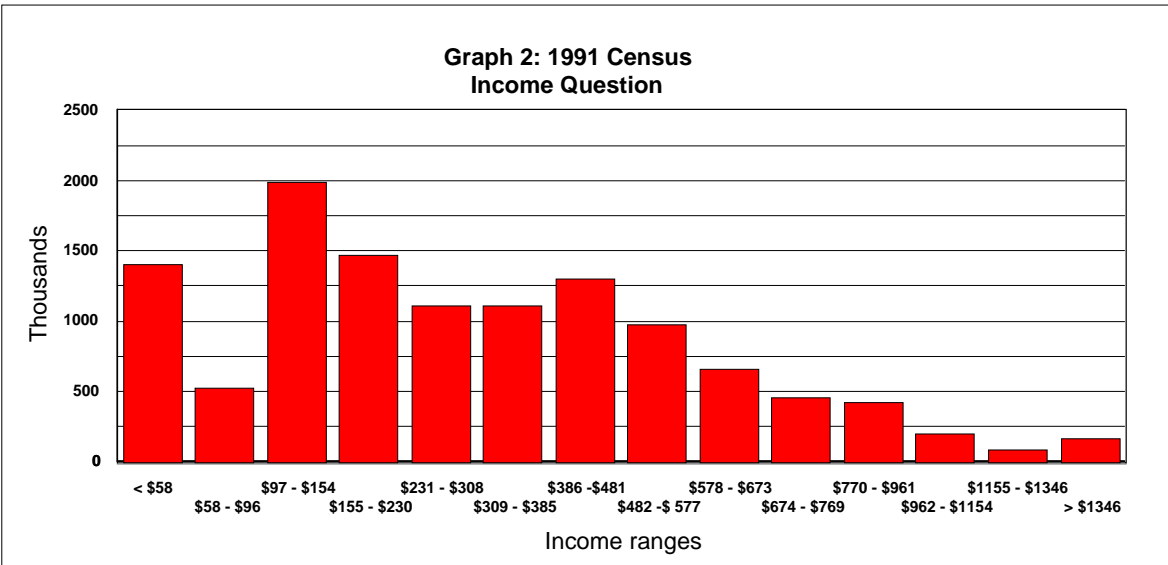
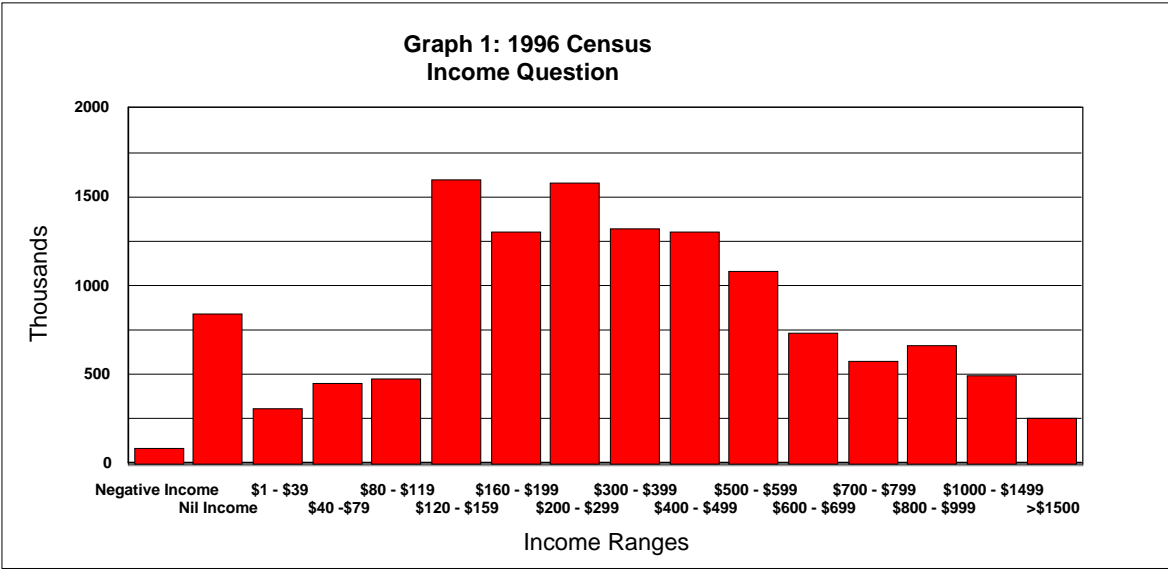
An accurate comparison of income data from the 1991 and 1996 Censuses is not possible for the following reasons:

- the income ranges for the two censuses are different, even after adjusting the 1991 Census income ranges for inflation. See Appendix 3; and
- the number of income response categories was increased from 14 in the 1991 Census to 16 in the 1996 Census, with the inclusion of the 'negative' and 'nil' income response options.

Consequently, the data analysis of both censuses focuses on the characteristics and trends of the data distribution rather than a direct comparison of the data therefore, the

data in Graphs 1 and 2 has not been adjusted for inflation. Graphs 1 and 2 show the income distribution from the 1996 and 1991 Censuses respectively. They also show that the recategorization of the four low income response options has reshaped the low income end of the histogram. This change allowed for relatively better data capture at the lower end of the income range. The lack of explicit 'nil' and 'negative' income response categories in the 1991 Census appears to have caused confusion and a high non-response rate amongst persons in these categories.

At the time, the data quality issues relating to the 1991 Census income question were thought to be associated with an apparent perception, by respondents, that the question was intrusive. However, as will be discussed in the non-response section for the 1996 Census (see section 2.3), this proposition is now considered to be an oversimplification and the focus of the analysis has shifted to the relative complexity of asking a single income question, and the incorrect perception by some respondents (in particular those who were not in the labour force or who were unemployed), that the income question did not apply to them.



The trend observed in Graphs 1 and 2 for the ‘middle’ income ranges is similar for both censuses. For the 1996 Census, data for respondents earning between \$120 and \$999 per week show a relatively similar pattern to that for income earners reporting an income between \$97 and \$961 per week in the 1991 Census. A better response rate and ‘real’ income growth might explain the skew to the right for ‘middle’ income earners in the 1996 Census income data compared to the 1991 Census data.

As seen in Graph 1 a large percentage of respondents, 5.3% of the population or 744,509 persons (see Appendix 4), reported earning a high income (above \$999 per week) in the 1996 Census. This was larger than the 3.4% of the population or 439,585 persons, (see Appendix 4) who reported earning a high income (above \$961 per week) in the 1991 Census. Likely explanations for this 55.9% increase in the number of respondents reporting high incomes in the 1996 Census are:

- an increase in 'real' income over the period 1991 to 1996, or
- the change in the order of the response options in the question (in descending income order in 1996). This means that respondents who marked the first available response option in the income question, would in fact have been reporting a high income. This differs from the situation in 1991 when the response options for income were in ascending order. The tendency by some respondents to select the first response option to a question without thinking about their response, is often referred to as ‘the list effect’.

2.3 *Non-Response Analysis*

The non-response rate (NRR) for the income question is usually higher than the NRR for most other questions. A likely explanation for this is the self-enumerated nature of the census, where respondents need to understand a single income question and to answer it correctly. Respondents have little opportunity to ask their collector for clarification of the question and what they should and should not include as income, to enable them to provide a complete and accurate response to the question.

During the 1996 Census, more calls to the Census Inquiry Service related to the income question than to any other question (4.6% of total calls were categorised as relating to the income question). Despite this the NRR for the income question in 1996 was 6.2%, which represented a fall of 33% (or -3.1% points) from the 1991 NRR of 9.3%. A comparison between the NRR by sex for the last three censuses is shown in Table 1.

Table 1: Income Non-Response Rates by Sex; Australia 1986, 1991 and 1996 Censuses

<i>Census</i>	<i>1986</i>	<i>1991</i>	<i>1996</i>
	<i>%</i>	<i>%</i>	<i>%</i>
Males	5.7	7.4	6.0
Females	7.7	11.1	6.4
Persons	6.7	9.3	6.2

Various alternatives for the income question in terms of wording, response options and position of the question on the form have been used in the last three censuses. However, the patterns of non-response shown in table 1 above suggest that the inclusion of a ‘nil’ income response option, in both the 1986 and 1996 Censuses, had a

positive impact on the response rate. In 1991 when such an option was omitted the total NRR rose from 6.7% to 9.3% for all respondents. Other changes to the question in 1996, such as the inclusion of a 'negative' income response option and the change in the order of income categories, are also thought to have contributed to the drop in the NRR to 6.2% in 1996.

In 1986, 1991 and 1996 the NRRs for females were higher than those for males (Table 1). However, in 1996 the gap between the rates for males and females was reduced to only 0.4 percentage points. This compares with differences of 2.0 and 3.7 percentage points in 1986 and 1991 respectively. It seems likely that the inclusion of the 'nil' income category was an important factor in reducing the overall NRR in 1996, particularly for females (a 42.3% decrease in NRR, from 11.1% in the 1991 Census to 6.4% in the 1996 Census) as they are more likely to report 'nil' income. Previous research suggests that persons without a job (for example, females staying at home performing unpaid work) may fail to respond to the income question because they mistakenly assume it only relates to working persons receiving an income (see Census Working Paper 96/1: *1991 Census Data Quality: Income*).

Table 2: Persons with Not Stated Income by Labour Force Status (LFS) and by Sex, Australia 1996 Census

<i>LFS Broad Group</i>	<i>Total males</i>	<i>% Males</i>	<i>Total females</i>	<i>% Females</i>	<i>Total all</i>	<i>% Total</i>
Employed	87,661	21.4	65,448	14.4	153,109	17.7
Unemployed	17,139	4.2	11,771	2.6	28,910	3.4
NILF	147,607	36.1	251,967	55.4	399,574	46.2
Not Stated (LFS)	156,685	38.3	125,767	27.6	282,452	32.7
Total Non-Response	409,092	100.0	454,953	100.0	864,045	100.0

'Employed' comprises: employee, employer, own account worker and contributing family worker; 'Unemployed' comprises: unemployed looking for full-time job and unemployed looking for part-time job; and 'Not in the labour force' (NILF) includes: pensioners, persons permanently unable to work, persons in institutions, persons not looking for work and students. (See *1996 Census Dictionary*, ABS Cat. No. 2901.0).

The table above shows that for the 1996 Census 46.2% of the total non-response to the income question was for persons who stated that they were 'not in the labour force' (NILF). A major contributor to this figure were females with 55.4% of total female non-response, compared with 36.1% of total male non-response. These figures would tend to indicate that this group did not believe that the income question applied to them, possibly because they did not regard their government transfer payments, such as pensions and allowances, to be income. Table 2 also shows that 32.7% of the non-responses for the income question were also non-responses for LFS. Of the total non-respondents to the income question only 17.7% were employed persons.

Table 3: Total Population Over 15 Years of Age by LFS, by Income Stated and by Income Not Stated; Australia 1996 Census

<i>LFS Broad Group</i>	<i>Income Stated</i>	<i>%</i>	<i>Income Not Stated</i>	<i>%</i>	<i>Total</i>
Employed	7,483,210	98.0	153,109	2.0	7,636,319
Unemployed	743,060	96.2	28,910	3.8	771,970
NILF	4,774,607	92.3	399,574	7.7	5,174,181
Not Stated (LFS)	49,975	15.1	282,452	84.9	332,427
Total	13,050,852		864,045		13,914,897

In the past it has been suggested that failure to answer the income question was due to its perceived intrusiveness, particularly by 'middle' and 'high' income earners. As Table 3 shows 98% of employed respondents answered the income question. This high response rate from this group suggests that 'middle' and 'high' income earners do not perceive the question as being intrusive. Table 3 indicates that the problem lies with those respondents who do not think that the income question applies to them as they do not earn an income. This group of persons are primarily those respondents who stated that they were NILF (399,574 persons) and those respondents who did not answer either the LFS question or the income question (282,452 persons).

2.4 Effect of Dummy Records on the Non-Response Rate

In 1996, dummy census forms were created by census collectors for dwellings from which no completed forms could be obtained. Collectors were instructed to record the total numbers of males and females for these dwellings if this was known. They were also instructed to complete any known details for the age, sex and marital status questions. Where collectors were unable to provide any details about the occupants of the dwelling and the dwelling did not appear to be unoccupied, the average number of males and females in non-dummy private dwellings in the same Collection District (CD) was used as a basis for imputing the number of persons present for dummy forms. Where necessary age, marital status and State of usual residence were also imputed. All other fields were set to 'Not stated' or 'Not applicable'. Where the collector believed the dwelling was unoccupied, he or she was instructed to mark this on the form and no imputation was undertaken for that dwelling.

In the 1996 Census, 190,722 persons were enumerated on dummy forms. This represents 1.4% of the total census person count. Historically, dummy records have been shown to account for up to 50% of the not stated rate for certain variables. The impact of dummy records on the 1996 Census income non-response rate is shown in Table 4 below. When dummy records are excluded, the total non-response rate is reduced by 1.3 percentage points from 6.2% to 4.9%.

Table 4: Effect of Dummy Records on Income Non-Response Rates, Australia 1996 Census

<i>Response</i>	<i>Count</i>	<i>%</i>	<i>Counts excluding dummy records</i>	<i>%</i>
Stated	13,050,852	93.8	13,050,852	95.1
Not Stated				
Dummy record	190,722	1.4	--	--
Not a Dummy record	673,323	4.8	673,323	4.9
<i>Total Not stated</i>	864,045	6.2	673,323	4.9
Total persons aged 15 and over	13,914,897	100.0	13,724,175	100.0

3. NEGATIVE AND NIL INCOME, 1996 CENSUS

3.1 Negative Income

In 1996 for the first time, the income question included a 'negative' income response category. The reason for including a 'negative' income response category was twofold. Firstly, it was in line with the ABS income standard which was developed after the 1991 Census. Secondly, it was expected that its inclusion would help to reduce the non-response rate which, as previously discussed (see section 2.3), peaked in 1991.

As part of the ABS program of producing statistical standards, an income standard was developed in August 1995 to ensure that income data were collected consistently across all ABS collections. The use of such a standard helped improve data quality and comparability between collections. The ABS income standard provides four alternative question modules, one of which is a single question module designed for inclusion on self-enumerated questionnaires. A version of this was used in the 1996 Census. For a detailed description of the four alternative income questions, refer to *ABS Standards for Cash Income Statistics* (ABS Cat. No. 1287.0).

The 'negative' income category was intended for respondents who were self-employed, or in a business or rental property situation where expenses were greater than income, and where this loss was greater than the total of all other income, benefits and allowances received. However, there were no instructions to this effect on either the census form itself or in the accompanying information booklet. Furthermore, focus group testing suggests that this response category was misunderstood by some persons who believed it applied to anybody living off their savings. It is also suspected that some people who supplemented their government (or private) transfer payments such as, unemployment benefits, pensions, gifts, and money from charities with money from their savings, did not consider these transfer payments to be income and consequently might have reported 'negative' income.¹

By definition 'negative' income should only have been reported by the owners of unincorporated businesses (the self-employed, own account workers, employers and partners). This response option should not have been used by employed persons or any other labour force categories.

¹ See Appendix 2, for the list of recommended income sources in the question. Also refer to the ABS publication, *Standards for Cash Income Statistics* (ABS Cat. No. 1287.0).

3.2 Characteristics of Persons Reporting 'Negative' Income

Table 5: Characteristics of Persons Reporting 'Negative' Income Values, Australia 1996 Census

<i>Labour force status:</i>	<i>Employee</i>	<i>Employer</i>	<i>Own account worker</i>	<i>Contributing family worker</i>	<i>Unemployed</i>	<i>NILF</i>	<i>Not stated</i>	<i>Total</i>	<i>Total count</i>
	%	%	%	%	%	%	%	%	
<i>Age</i>	<i>Persons</i>	<i>Persons</i>	<i>Persons</i>	<i>Persons</i>	<i>Persons</i>	<i>Persons</i>	<i>Persons</i>	<i>Persons</i>	<i>Persons</i>
15 - 19	0.8	0.0	0.0	0.2	0.6	16.2	0.2	19.0	15,369
20 - 64	12.4	4.3	11.9	3.0	6.7	32.0	1.2	71.6	57,751
> 64	0.6	0.3	0.8	0.3	0.1	7.4	n.a.	9.4	7,685
Total	13.8	4.6	12.7	3.5	8.4	55.6	1.4	100.0	80,805

n.a.: refers to values smaller than 0.01%.

Table 5 shows that in the 1996 Census 80,805 persons aged 15 years or more, stated that they had a 'negative' income. Significantly, 55.6% of the total response for the 'negative' income option were persons NILF. Of these 16.2% were between 15 and 19 years of age. A large proportion of these persons are likely to have mistakenly used the negative income response category. For example, a proportion of persons in this group might have considered their Higher Education Contribution Scheme accruing debt to be a current expense and consequently chosen the negative response option, when in fact they should have chosen either the 'nil' income or one of the low income categories. Also, 32.0% of respondents who reported a 'negative' income, reported that they were NILF and were between the ages of 20 and 64 years of age. Persons aged over 64 accounted for 9.4% of those reporting 'negative' income.

13.8% of respondents reporting negative income were in the LFS classification of 'Employee'. Of these 12.4% were between 20 and 64 years of age. There was apparently a misunderstanding of this response option as wage and salary earners should have a positive income value. Two possible explanations for this response rate are that wage and salary earners who do not earn enough to cover their living costs and had to draw on their savings perceived this negative cash flow to be 'negative' income. Also, farmers who ran their businesses at a loss in a particular year and took a part-time job but were still unable to cover their farm expenses therefore recording an overall loss, reported this as 'negative' income on their Census form. While this would be the correct answer for income in this situation, it may or may not be the correct answer for LFS depending on whether the person spent more hours working at their paid employment or at their business.

Another LFS classification contributing to total negative income was the 'unemployed', who made up 8.4% of the total. As explained in section 2.3, many respondents in this classification did not appear to consider government transfer payments such as unemployment benefits, to be income. Consequently, they perceived their total expenses to be higher than their income and therefore reported a 'negative' income. Such people may also have been drawing on their savings to supplement their transfer payments and considered this to be 'negative' income.

Finally, 17.3% of the negative income response was reported by those who would have been expected to report a negative income, namely employers, 4.6% and own account workers, 12.7%.

3.3 Characteristics of Persons Reporting 'Nil' Income

A 'nil' income option was added as a response option for the 1996 Census, with the similar aim as for the negative income response option, that is, to reduce non response by providing an explicit response option appropriate to persons who were not employed and who, for some reason were not eligible for transfer payments. This response option also aimed to capture that proportion of the labour force that was considered to be unpaid workers (persons who stayed at home and did not receive any form of income).

Table 6: Characteristics of Persons Reporting Nil Income Values; Australia 1996 Census

<i>Labour force status:</i>	<i>Employee</i>	<i>Employer</i>	<i>Own account worker</i>	<i>Contributing family worker</i>	<i>Unemployed looking for work</i>	<i>NILF</i>	<i>Not stated</i>	<i>Total</i>	<i>Total counts</i>
	%	%	%	%	%	%	%	%	
<i>Age</i>	<i>Persons</i>	<i>Persons</i>	<i>Persons</i>	<i>Persons</i>	<i>Persons</i>	<i>Persons</i>	<i>Persons</i>	<i>Persons</i>	<i>Persons</i>
15 - 19	0.4	0.2	0.0	0.2	3.4	36.6	0.1	40.7	342,959
20 - 64	2.1	0.2	0.9	1.4	7.5	42.7	0.8	55.4	467,749
> 64	0.0	0.0	0.0	0.0	0.0	3.7	0.0	3.9	32,951
Total (a)	2.5	0.4	0.9	1.6	10.9	83.0	0.9	100.0	843,659

(a) Some totals do not add up due to rounding.

There is a higher response to the 'nil' income than to the 'negative' income response option. As Table 6 shows, a total of 843,659 persons or 6.1% of the total population (refer to Graph 1 and Appendix 4), selected the nil income response option as opposed to 80,805 (Table 5) or 0.6% of the total population who selected the negative income response option. This would be expected as unprofitable, unincorporated businesses do not stay in operation for long.

As seen in Table 6, 83.0% of 'nil' income respondents reported that they were not in the labour force. A large contributor to this figure (36.6%) were respondents between the ages of 15 and 19 years, many of whom were likely to be still in full-time education. Also, 42.7% were persons between the age of 20 and 64 years who reported that they were NILF and had a 'nil' income. A possible explanation is that these people were either not receiving any form of income or misinterpreted this response option and therefore did not include any transfer payments received as income, and consequently reported a 'nil' income.

In conclusion, it appears that many respondents who reported that they were not in the labour force and answered the income question, did not consider either government or private transfer payments such as benefits, family allowances and student allowances to be income, consequently they directed their answers to either the 'nil' or 'negative' income response options. However, it is not possible to determine which respondents accurately reported 'negative' or 'nil' income, and which respondents misinterpreted the response options.

4. COMPARISON OF THE 1996 CENSUS WITH THE 1995-96 SURVEY OF INCOME AND HOUSING COSTS (SIHC), AND TAXATION STATISTICS 1995-96

This section compares the response distributions for income from the 1996 Census with the 1995-96 Survey of Income and Housing Costs (SIHC), and Taxation Statistics 1995-96.

4.1 Comparison of Distributions of Income from the 1996 Census with SIHC

Although the census and SIHC both collect data on income, they are not strictly comparable due to differences in scope, coverage, timing and collection methodology. Refer to Census Working Paper 96/1: *1991 Census Data Quality: Income*, for a more detailed explanation of the differences between the census and survey methodologies.

Factors contributing to differences in the data include under-enumeration in the census for which census income estimates have not been adjusted, the use in the SIHC of benchmarks based on earlier census and population projections (rather than the current census population), the inclusion of permanent defence force personnel in census estimates, differing methods of adjustment for non-response in the survey and the census, the personal interview approach adopted in the survey as opposed to self-enumeration in the census and the more detailed questions used in SIHC to determine income.

The distribution of income for the 1996 Census was relatively similar to the 1995-96 SIHC, although differences were noted at the lower end of the income distribution range. For the '\$1 - \$39 per week' response category, census figures were 43.6% lower. For the 'negative' income response option, census figures were 40% lower than the survey estimates and 20% lower in the \$40 - \$79 per week response category. This can be seen in Graph 3 and Table 7 which both compare the income distributions of the two collections. A possible explanation for these differences is the self-enumerated nature of the census. It has been shown that these 'low' and 'nil' or 'negative' income groups tend to misinterpret or not relate to the income question, which results in a higher non-response rate from low income earners in the census.

The differences at the middle and higher income ranges were small. The census had a greater proportion of persons in the categories between \$160 and \$399 but again a lower proportion in the higher income ranges, with the exception of the category \$1,500 or more, where the census had a higher response rate (20.0% more than the survey). As seen in Appendix 4, 1.2% of people reported that they were in the top income range, with an income greater than \$1346 per week in the 1991 Census. In the 1996 Census 1.8% of people reported that they were in the top income range with an income greater than \$1500 per week. That is, the percentage of the total population reporting the top income range in both the 1991 and 1996 Censuses is similar.

It is reasonable to assume that the figure from the 1995-96 SIHC with 1.5% of respondents reporting that they were in the top income range, (\$1500 or more) is the most likely of the three figures to be accurate, as it is an interviewer based collection. If this is the case and as the figure for the 1991 Census is lower (1.2%), it follows that in the 1991 Census 0.3% of the population understated their income. This may have been the result of the 'list effect' as the income ranges on the 1991 Census Form were listed from

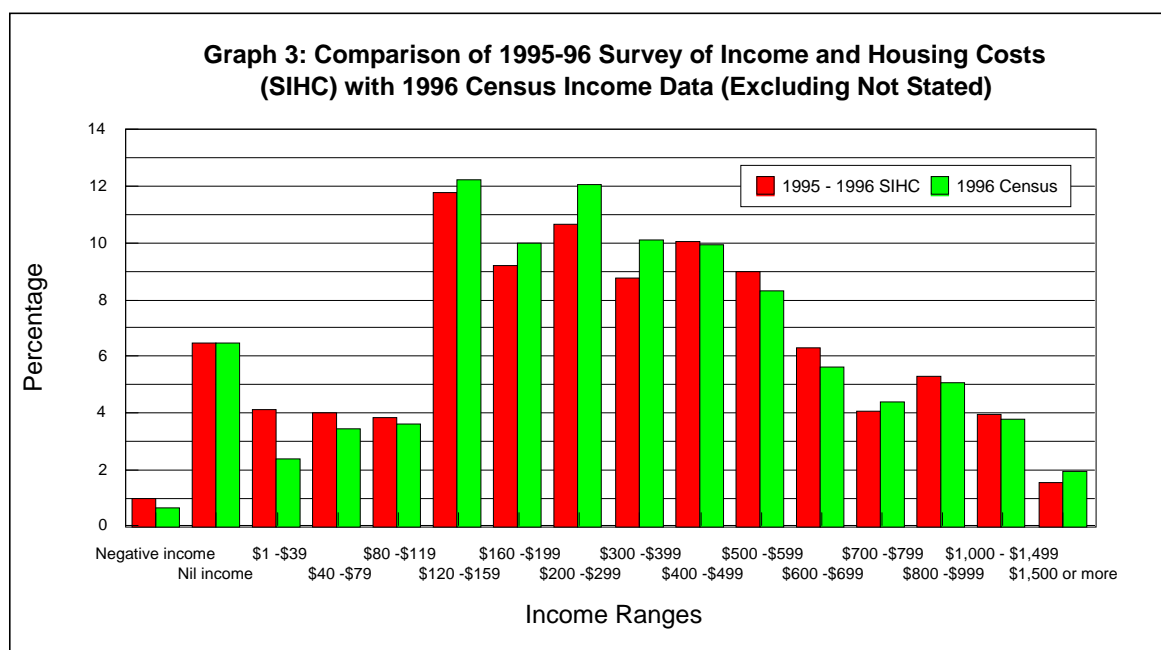
the lowest income range to the highest. In the 1996 Census 1.8% of the population reported that they were in the top income range. However, in 1996 the income ranges in the income question went from the highest to the lowest, with the top range being the first option rather than the last, as was the case on the 1991 Census Form. It is likely that most of the increase in people reporting the top income range in the 1996 Census was due to the 'list effect', rather than to people deliberately bragging and overstating their income.

The results are not surprising given the probability that persons with low or no income perceived the question as irrelevant to them, resulting in a high non-response rate for these groups in the census. Factors such as the self-enumerated nature of the census compared with the more detailed collection of data from the income survey may explain the small differences. Generally, at this level of aggregation there is little difference between the results. This can be seen in both Table 7 and Graph 3 below.

TABLE 7: Distribution of Income for Persons Aged 15 and Over, 1995-96 Survey of Income and Housing Costs (SIHC) and 1996 Census

<i>Income (per week)</i>	<i>1995-96 SIHC</i>	<i>1996 Census</i>	<i>% Difference</i>
	<i>%</i>	<i>%</i>	
Negative Income	1.0	0.6	-40.0
Nil Income	6.5	6.1	-6.2
\$1 - \$39	3.9	2.2	-43.6
\$40 - \$79	4.0	3.2	-20.0
\$80 - \$119	3.8	3.4	-10.5
\$120 - \$159	11.8	11.5	-2.5
\$160 - \$199	9.2	9.4	2.2
\$200 - \$299	10.7	11.3	5.6
\$300 - \$399	8.8	9.5	8.0
\$400 - \$499	10.1	9.3	-7.9
\$500 - \$599	9.0	7.8	-13.3
\$600 - \$699	6.3	5.3	-15.9
\$700 - \$799	4.1	4.1	0.0
\$800 - \$999	5.3	4.8	-9.4
\$1000 - \$1,499	4.0	3.5	-12.5
\$1500 or more	1.5	1.8	20.0
Not stated	0.0	6.2	n.a.
Total	100.0	100.0	--
Total Counts ('000s)	13,973	13,915	--

n.a.: The census records a code of "not stated" for respondents who could not be coded to an income category. The survey has no such category. Partial non-response in the survey is automatically adjusted for by the weighting process.



4.2 Comparison of Distributions of Income from the 1996 Census with Taxation Statistics 1995-96

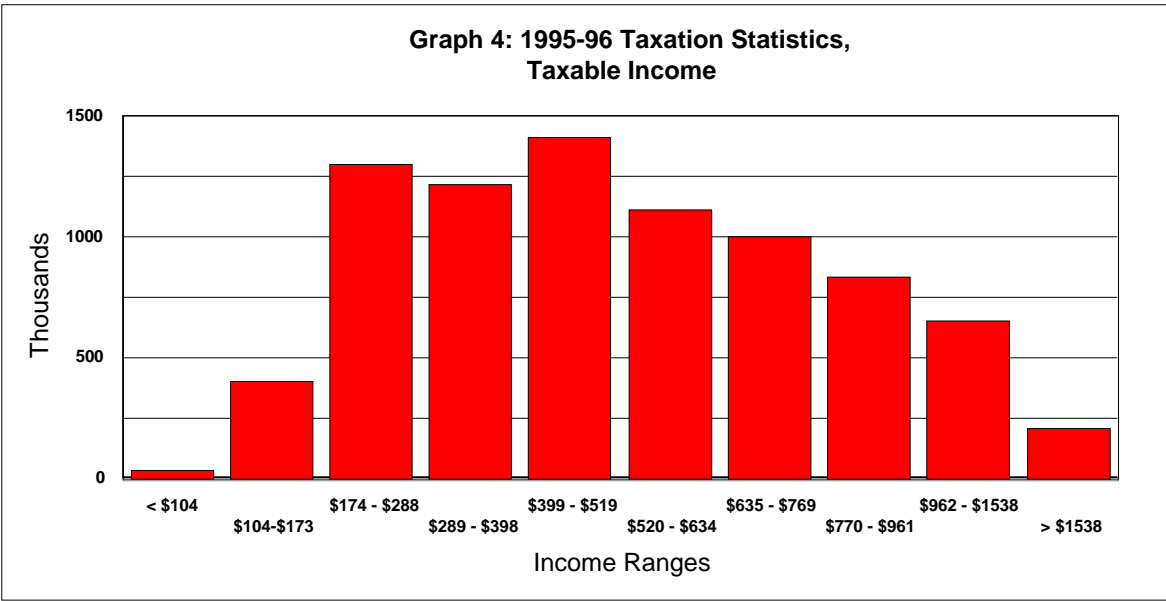
This section compares income data from the 1996 Census with income data collected by the Australian Taxation Office (ATO) for 1995-96. The following are the main methodological differences between the two sets of data:

- income tax data are based on taxable income not gross income as collected in the census;
- 1996 Census data have a broader coverage with a total population of 13,050,852 persons, while tax data cover 8,165,622 persons (generally excluding persons with income earnings below the minimum taxable income bracket);
- Census data capture transfer payments for persons aged 15 and above and for persons not in the labour force, whereas taxation data do not always capture these payments;
- different income ranges (see Appendix 5); and
- income tax data were collected for the financial year ended 30 June 1996 while the Census income data were collected on Census night, 6 August 1996, and asked for the gross income the respondent 'usually receives each week'.

For the purpose of this paper, Table 4 on page 30 of *1995-96 Taxation Statistics*, was compared with the 1996 Census income data.

The analysis focuses on the trends shown by both sets of data. Despite their methodological differences, the underlying trends are similar for both collections. A comparison of Graphs 1 and 4 shows that low income earners were captured better by the census. This is because taxation data were generally not collected for persons who earn below the income tax threshold of \$5,401 per year (\$104 per week) in 1996. By contrast the census aimed to capture income from every person over 15 years of age in Australia on Census night. There were 3,753,601 respondents who reported earning less

than \$8,320 per year (\$160 per week) in the 1996 Census (see Appendix 4), compared to 438,020 persons who reported a taxable income lower than \$9,000 per annum (\$173 per week) to the ATO for the year ended 30 June 1996 (see Appendix 5).



Note: Taxation income data are published on an annual basis. For the purpose of this paper they have been converted to weekly income.

However, for census respondents earning more than \$8,319 per year (\$159 per week) and ATO data providers earning more than \$8,999 per year (\$173 per week), both sets of data show a similar trend (refer to Graphs 1 and 4).

It can be concluded that both sets of data are relatively comparable, despite their methodological differences. The shape of the histograms for ‘middle’ and ‘high’ income earners for both the 1996 Census and the 1995-96 Taxation Statistics are similar. It is at the lower income end where data from the 1996 Census is more comprehensive and hence more useful for data users as many of this lower income group are out of scope for the 1995-96 Taxation Statistics.

5. THE 2001 CENSUS INCOME QUESTION

The Income question to be used in the 2001 Census is basically the same as the question used in the 1996 Census. Only very minor changes will be made. These are:

- changes to the wording of two of the reference sources of income, from 'family payment' in 1996 to 'family allowance' in 2001 and from 'additional family payment' in 1996 to 'parenting allowance' in 2001; the wording was revised to make it consistent with current Centrelink terminology;
- the inclusion of a dot point at the end of the question to refer respondents to the Census Guide for more information;
- the inclusion in the Census Guide of an explanation of negative income; and
- the question was at the top of the page for the 1996 Census but will be at the bottom of the page for the 2001 Census.

Several factors were taken into consideration when making the decision to retain the income ranges for the 2001 Census the same as they were for the 1996 Census. In particular:

- the inflation rate, as measured by the change in Consumer Price Index (CPI) average of all groups, all capital cities, in the past few years has been relatively low:
Change between CPI March 1996 and CPI March 1997 = 1.2%;
Change between CPI March 1997 and CPI March 1998 = -0.2%;
Change between CPI March 1998 and CPI March 1999 = 1.2%;
Change between CPI March 1999 and CPI March 2000 = 2.8%; and
- maintaining the same income ranges will ensure the comparability of the results from the 2001 Census with the 1996 Census.

6. CONCLUSION

One of the most important data quality achievements for the income question in the 1996 Census was the decrease in the non-response rate from 9.3% in the 1991 Census to 6.2% in the 1996 Census (3.1 percentage points). When dummy records were excluded, the total non-response rate was further reduced from 6.2% to 5.0% (1.2 percentage points). ABS believes the following factors contributed to this drop in the non-response rate:

- the increase in the number of response options from fourteen in the 1991 Census to sixteen in the 1996 Census;
- the introduction of two new income response categories: 'Nil income' and 'Negative income', meaning that persons with no or low income had an appropriate option to which they could relate their financial situation;
- the repositioning of the income question at the top of the page in the 1996 Census form, from second from the top in the 1991 Census form; and
- the inclusion of more examples of income sources in the question.

However, despite the decrease in non-response, there are some remaining data quality concerns (as listed below) associated with space restrictions and the complexity of asking a single income question in a self-enumerated census form where no, or very little, assistance is available to the respondent. These are:-

- the relatively high non-response rates by persons who are likely to be low income earners. In particular persons who are not in the labour force and the unemployed;
- a general tendency by respondents not to state or to underestimate their income. However, as this paper shows this is more evident with low income earners and may be because they do not regard their government and private transfer payments as income;
- the difficulty for some respondents to accurately recollect all their sources of income and to include these in their response. This is despite including a comprehensive list of examples of income sources in the census income question; and
- The tendency by respondents to report only their main source of income (wages or salary) and to omit any additional income sources, such as interest received, dividends, overtime and government transfer payments.

Irrespective of the differences between the 1991 Census income question and the 1996 Census income question, it has been shown that the 1996 Census income question captured in a satisfactory way income distribution patterns as at 6 August 1996. This can be verified by comparing 1996 Census income data with both ABS 1995-96 Survey of Income and Housing Costs and Taxation Statistics 1995-96 data. A comparison of income data distribution patterns between the 1996 Census and these collections shows that despite the different collection methodologies, the reported income distributions were relatively similar. 1996 Census data captures income in the lower income ranges better than 1995-96 Taxation Statistics data while data for the middle and higher income ranges is consistent for both collections.

It can be concluded that, despite the limitations of collecting census information by self-enumeration, the data can still be used as a comparative indicator of income levels for small geographical areas and small population groups. Comparison of the 1996 Census data with data from the 1991 Census, 1995-96 SIHC and 1995-96 Taxation Statistics shows an overall similarity, confirming the validity of the data.

APPENDIX 1: 1991 Census Income Question

29 What is the gross income (including pension and allowances) that the person usually receives each week from all sources?

•Count all income including:

family allowance
family allowance supplement
pensions
unemployment benefits
student allowance
maintenance
(child support)
worker's compensation
superannuation
wages
salary
overtime
dividends
rents received
business or farm income
(less expenses of operation)
interest received.

• Do not deduct:

tax
superannuation
health insurance

- () Less than \$58 per week
(Less than \$3,001 per year)
- () \$58 to \$96 per week
(\$3,001 to \$5,000 per year)
- () \$97 to \$154 per week
(\$5,001 to \$8,000 per year)
- () \$155 to \$230 per week
(\$8,001 to \$12,000 per year)
- () \$231 to \$308 per week
(\$12,001 to \$16,000 per year)
- () \$309 to \$385 per week
(\$16,001 to \$20,000 per year)
- () \$386 to \$481 per week
(\$20,001 to \$25,000 per year)
- () \$482 to \$577 per week
(\$25,001 to \$30,000 per year)
- () \$578 to \$673 per week
(\$30,001 to \$35,000 per year)
- () \$674 to \$769 per week
(\$35,001 to \$40,000 per year)
- () \$770 to \$961 per week
(\$40,001 to \$50,000 per year)
- () \$962 to \$1,154 per week
(\$50,001 to \$60,000 per year)
- () \$1,155 to \$1,346 per week
(\$60,001 to \$70,000 per year)
- () More than \$1346 per week
(More than \$70,000 per year)

APPENDIX 2: 1996 Census Income Question

29 What is the gross income (including pensions and allowances) that the person usually receives each week from all sources?

• **Mark one box only.**

• **Count all income for each**

person including:

family payment

additional family payment

rental assistance

pensions

unemployment benefits

student allowance

maintenance

(child support)

worker's compensation

superannuation

wages

salary

overtime

commissions and bonuses

interest received

dividends

rents received

(less expenses of operation)

business or farm income

(less expenses of operation)

• **Do not deduct:**

tax

superannuation

health insurance

- () \$1,500 or more per week
(\$78, 000 or more per year)
- () \$1000 - \$1,499 per week
(\$52, 000 - \$77,999 per year)
- () \$800 - \$999 per week
(\$41,600 - \$51,999 per year)
- () \$700 - \$799 per week
(\$36,400 - \$41,599 per year)
- () \$600 - \$699 per week
(\$31,200 - \$36,999 per year)
- () \$500 - \$599 per week
(\$26,000 - \$31,199 per year)
- () \$400 - \$499 per week
(\$20, 800 - \$25, 999 per year)
- () \$300 - \$399 per week
(\$15,600 - \$20,799 per year)
- () \$200 - \$299 per week
(\$10,400 - \$15,599 per year)
- () \$160 - \$199 per week
(\$8,320 - \$10,399 per year)
- () \$120 - \$159 per week
(\$6,240 - \$8,319 per year)
- () \$80 - \$119 per week
(\$4,160 - \$6,239 per year)
- () \$40 - \$79 per week
(\$2,080 - \$4159 per year)
- () \$1 - \$39 per week
(\$1 - \$2,079 per year)
- () Nil income
- () Negative income

APPENDIX 3: 1991 and 1996 Censuses Income Ranges

<i>1991 Income Ranges</i>	<i>1991 Income Ranges Adjusted for Inflation (1996 dollars)</i>	<i>Best Matching 1996 Income Ranges</i>
Less than \$58 per week	Less than \$66 per week	Negative Income/ Nil Income/\$1 to \$39/\$40 to \$79 per week
\$58 to \$96 per week	\$66 to \$108 per week	\$80 to \$119 per week
\$97 to \$154 per week	\$109 to \$174 per week	\$120 to \$159 per week
\$155 to \$230 per week	\$175 to \$260 per week	\$200 to \$299 per week
\$231 to \$308 per week	\$261 to \$349 per week	\$300 to \$399 per week
\$309 to \$385 per week	\$350 to \$435 per week	\$300 to \$399 per week
\$386 to \$481 per week	\$436 to \$544 per week	\$400 to \$499 per week
\$482 to \$577 per week	\$545 to \$652 per week	\$500 to \$599 per week
\$578 to \$673 per week	\$653 to \$760 per week	\$700 to \$799 per week
\$674 to \$769 per week	\$761 to \$869 per week	\$800 to \$999 per week
\$770 to \$961 per week	\$870 to \$1,086 per week	\$800 to \$899 per week
\$962 to \$1,154 per week	\$1,087 to \$1,304 per week	\$1,000 to \$1,499 per week
\$1,155 to \$1,346 per week	\$1,305 to \$1,521 per week	\$1,000 to \$1,499 per week
More than \$1,346 per week	More than \$1,521 per week	More than \$1,500 per week

<i>1991 Income Ranges</i>	<i>1996 Income Ranges</i>
Less than \$58 per week	Negative income Nil income \$1 to \$39 per week \$40 to \$79 per week
\$58 to \$96 per week	\$80 to \$119 per week
\$97 to \$154 per week	\$120 to \$159 per week
\$155 to \$230 per week	\$160 to \$199 per week \$200 to \$299 per week
\$231 to \$308 per week \$309 to \$385 per week	\$300 to \$399 per week
\$386 to \$481 per week \$482 to \$577 per week \$578 to \$673 per week	\$400 to \$499 per week \$500 to \$599 per week \$600 to \$699 per week \$700 to \$799 per week
\$674 to \$769 per week \$770 to \$961 per week	\$800 to \$999 per week
\$962 to \$1,154 per week \$1,155 to \$1,346 per week	\$1,000 to \$1,499 per week
More than \$1,346 per week	More than \$1,500

APPENDIX 4: Distribution of Income, Persons Aged 15 and Over, 1991 and 1996 Censuses

Distribution of Income, Persons Aged 15 and Over, 1996 Census

<i>Income per week</i>	<i>Population numbers</i>	<i>%</i>
Less than 0	80,805	0.6
0	843,959	6.1
\$1 to \$39	309,687	2.2
\$40 to \$79	448,567	3.2
\$80 to \$119	473,500	3.4
\$120 to \$159	1,597,083	11.5
\$160 to \$199	1,302,223	9.4
\$200 to \$299	1,575,312	11.3
\$300 to \$399	1,318,332	9.5
\$400 to \$499	1,299,394	9.3
\$500 to \$599	1,083,441	7.8
\$600 to \$699	737,371	5.3
\$700 to \$799	572,994	4.1
\$800 to \$999	663,675	4.8
\$1,000 to \$1,499	494,283	3.5
More than \$1,500	250,226	1.8
Not Stated	864,045	6.2
Total	13,914,897	100.0

Distribution of Income, Persons Aged 15 and Over, 1991 Census

<i>Income per week</i>	<i>Income per week (adjusted for inflation to 1996 dollars)</i>	<i>Population numbers</i>	<i>%</i>
Less than \$58	Less than \$65	1,399,581	10.7
\$58 to \$96	\$65 to \$108	527,882	4.0
\$97 to \$154	\$109 to \$174	1,988,000	15.2
\$155 to \$230	\$175 to \$259	1,470,731	11.2
\$231 to \$308	\$260 to \$347	1,103,942	8.4
\$309 to \$385	\$348 to \$434	1,113,749	8.5
\$386 to \$481	\$435 to \$542	1,296,549	9.9
\$482 to \$577	\$543 to \$650	973,027	7.4
\$578 to \$673	\$651 to \$758	654,129	5.0
\$674 to \$769	\$760 to \$867	460,763	3.5
\$770 to \$961	\$868 to \$1,083	425,656	3.3
\$962 to \$1,154	\$1,084 to \$1,301	193,964	1.5
\$1,155 to \$1,346	\$1,302 to \$1,517	84,703	0.7
More than \$1,346	More than \$1,517	160,918	1.2
Not Stated		1,229,843	9.4
Total		13,083,437	100.0

APPENDIX 5: Taxation Statistics 1995-96, Taxable Individuals by Fine Grade of Taxable Income, Australia

<i>Grade of Taxable Annual Income</i>	<i>Weekly Equivalent</i>	<i>Total number</i>	<i>%</i>
Under \$5,401	Under \$103	33,843	0.4
\$5,401 - \$6,999	\$104 - \$134	110,477	1.4
\$7,000 - \$8,999	\$135 - \$172	293,700	3.6
\$9,000 - \$10,999	\$173 - \$211	413,107	5.1
\$11,000 - \$12,999	\$212 - \$249	451,182	5.5
\$13,000 - \$14,999	\$250 - \$288	433,927	5.3
\$15,000 - \$16,999	\$289 - \$326	418,890	5.1
\$17,000 - \$18,999	\$327 - \$365	412,078	5.0
\$19,000 - \$20,700	\$366 - \$398	387,353	4.7
\$20,701 - \$22,999	\$399 - \$442	526,611	6.4
\$23,000 - \$24,999	\$443 - \$480	448,592	5.5
\$25,000 - \$26,999	\$481 - \$519	434,561	5.3
\$27,000 - \$28,999	\$520 - \$557	401,347	4.9
\$29,000 - \$30,999	\$558 - \$596	369,864	4.5
\$31,000 - \$32,999	\$597 - \$634	340,847	4.2
\$33,000 - \$34,999	\$635 - \$673	311,075	3.8
\$35,000 - \$38,000	\$674 - \$730	437,032	5.4
\$38,001 - \$39,999	\$731 - \$769	252,286	3.1
\$40,000 - \$44,999	\$770 - \$865	496,979	6.1
\$45,000 - \$50,000	\$866 - \$961	334,854	4.1
\$50,001 - \$59,999	\$962 - \$1,153	373,323	4.6
\$60,000 - \$69,999	\$1,154 - \$1,346	182,954	2.2
\$70,000 - \$79,999	\$1,347 - \$1,538	97,735	1.2
\$80,000 - \$89,999	\$1,539 - \$1,730	57,464	0.7
\$90,000 - \$99,999	\$1,731 - \$1,923	35,512	0.4
\$100,000 - \$199,999	\$1,924 - \$3,846	88,768	1.1
\$200,000 - \$299,999	\$3,847 - \$5,769	12,589	0.2
\$300,000 - \$399,999	\$5,770 - \$7,692	4,075	0.1
\$400,000 - \$499,999	\$7,693 - \$9,615	1,714	0.0
\$500,000 - \$999,999	\$9,616 - \$19,230	2,210	0.1
\$1,000,000 and Over	\$19,231 and Over	673	0.0
Total		8,165,622	100.0

0.0¹: Values rounded to zero

Reference List

International Forum for Census Dissemination, Diffusion no. 13, June 1996: *Asking Income in the Australian Census*.

Australian Bureau of Statistics (1996) Census Working Paper 96/2: *Census Form Design Testing Program*.

Australian Bureau of Statistics (1996) Census Working Paper 96/1: *1991 Census Data Quality: Income*.

Australian Bureau of Statistics (August 1996) *Labour Force, Australia Preliminary*. Cat. No. 6202.0.

Australian Bureau of Statistics (1997) *Standards of Cash Income Statistics*. Cat. No. 1287.0.

Australian Taxation Office (1996) *Taxation Statistics 1995-96, A Summary of Taxation, Superannuation and Child Support Statistics*.

Census Working Papers

- 96/1 *1991 Census Data Quality: Income*
- 96/2 *1996 Census Form Design Testing Program*
- 96/3 *1996 Census of Population and Housing: Digital Geography Technical Information Paper*
- 97/1 *1996 Census: Homeless Enumeration Strategy*
- 99/1 *1996 Census: Industry Data Comparison*
- 99/2 *1996 Census: Labour Force Status*
- 99/3 *1996 Census Data Quality: Housing*
- 99/4 *1996 Census: Review of Enumeration of Indigenous Peoples in the 1996 Census*
- 99/5 *2001 Census: Indigenous Enumeration Strategy Draft*
- 99/6 *1996 Census Data Quality: Occupation*
- 00/1 *1996 Census Data Quality: Journey to Work*
- 00/2 *1996 Census Data Quality: Qualification Level and Field of Study*
- 00/3 *1996 Census Data Quality: Industry*

If you would like a copy of any of these papers, or have any other queries, please contact Rosa Gibbs on (02) 6252 5942 or Email: rosa.gibbs@abs.gov.au

The papers are also available on the ABS website at www.abs.gov.au