Australia 2001: Survey Information

Summary table

Generic information	
Name of survey	Survey of Income and Housing Costs (SICH)
Institution responsible	Australian Bureau of Labour Statistics (ABS)
Frequency	Continuous from 1994-95 to 1997-98, and then in 99-00, 2000-01 and 2002-03
Survey year / Wave	2000-01
Collection period	July 2000 to June 2001
Survey structure	Cross-sectional
Coverage	Individuals and families resident in private dwellings throughout Australia
Geographic information	
Files delivered	Basic Confidentialised Unit Record File (CURF), including a household, an income unit and a persons' file.
Sample size	
Households	Final sample of 6,786 households out of an initial sample of approximately 7,000 dwellings
Individuals	Final sample of 13,193 persons 15 years old and over out of an initial sample of approximately 15,500 persons over the age 15
Sampling	
Sampling design	Sub-sample of private dwellings included in the ABS Monthly Population Survey (MPS); the MPS sample is a multistage selection of private dwellings and a list sample of other dwellings
Sampling frame	
Questionnaires	There is one personal questionnaire for all persons aged 15 and over, and which collects data on income and housing. Information about personal and household characteristics, including part of the labour force information, is collected as part of the LFS, and then later combined with the income, housing and other info from the SIHC interview.
Standard classifications	
Education	Only distinguishes persons without qualifications from those with higher/bachelor degree / postgraduate diploma or other post-school qualifications
Occupation	National classification (9 categories)
Industry	National classification (17 categories)
Income	
Reference period	Income in SIHC is collected for two reference periods, the financial year preceding the date of interview (July 1999-June 2000) and a much shorter one centered around interview time, i.e. at, or close to, the time of interview (current income).
Unit of collection	All income sources are collected for each individual aged 15 or over.
Period of collection	Income is collected using a number of different reporting periods, such as the last financial year for own business and property income, and the usual payment for a period close to time of interview for wages and salaries, other sources of private income and government cash transfers.
Gross/net	Income as collected in the survey is gross income before the deduction of personal income tax and Medicare levy. Data on income tax and Medicare levy paid are not collected in the survey, but modelled by ABS.
Data editing / processing	
Consistency checks	A combination of clerical and computer-based systems is used to ensure completeness and consistency of the questionnaires and to check that logical sequences have been followed; nothing has been incorrectly included or excluded on the questionnaire; all the necessary items are present; and that specific values lie within valid ranges.

Weighting	Final weights are calculated through an iterative procedure in which initial weights (inverse of the probability of selection) are adjusted by a calibration process to ensure that survey estimates conform to independently estimated benchmarks (numbers of persons aged 15 and over, numbers of children under age 15, numbers of households, and the value of government benefit cash transfers). Weights inflate to total population.
Imputation	Extensive imputation

This document draws extensively from "Survey of Income and Housing Costs – User Guide", published in 1997 by the ABS and "2000-01 Survey of Income and Housing Costs Australia – Confidentialised Unit Record File (CURF) Technical Paper", published in August 2003 by ABS.

Table of contents:

A. General characteristics

Official name of the survey/data source:
Survey of Income and Housing Costs / SIHC

Administrative Unit responsible for the survey:

Living Conditions Section
Australian Bureau of Statistics (ABS)
Locked Bag 10 – Belconnen – ACT - 2616
Web: www.abs.gov.au

The SIHC was conducted continuously from 1994–95 to 1997–98, and then in 1999–2000, 2000–01 and 2002–03. The results from the 2002–03 SIHC which included an expanded sample of 11,000 households (up from about 7,000 households in earlier years), were released in 2004. From 2003–04 the income component of the former Household Expenditure Survey (HES) has been expanded in the new, six-yearly Household Income and Expenditure Survey (HIES), an 11,000 household survey which also incorporates a number of other changes to improve income estimation and analysis. In between the six-yearly HIES cycles, there will be two cycles of an 11,000 household SIHC (to be conducted next in respect of each of 2005–06 and 2007–08), which together with the HIES provide an ongoing biennial household income survey.

Previous surveys of income were conducted by the Australian Bureau of Statistics (ABS) in 1979, 1982, 1986 and 1990. These surveys were generally conducted over a two-month period, compared to a twelve-month period for the SIHC. Compared with income surveys conducted previously, the SIHC also included improvements to the survey weighting and estimation procedures, changes to the population in scope and changes to interviewing methods.

The Survey of Income and Housing Costs (SIHC) collects information on sources of income, amounts received and characteristics of persons aged 15 years and over resident in private dwellings throughout non-sparsely settled areas of Australia.

Some specific examples of the recent uses of the income survey results include:

- identification of population groups with low incomes and their characteristics;
- assessing the adequacy of government pension and allowance rates;
- estimation of the effects of changes to eligibility threshold income levels for government cash benefits e.g. in terms of the number of families affected; and
- examination of the impact of tax changes on the income of individuals and families with differing characteristics.

Some specific examples of use made of housing data from the survey include:

- identification of characteristics of families with housing costs in relation to their income;
- estimation of the number of low income families who need assistance;
- estimation of the level of assistance (e.g. rent subsidy and other forms of assistance) required to enable those on low incomes to meet their housing needs; and
- comparison of the total housing costs of the different tenure types e.g. renting versus buying.

B. Population, sampling size and sampling methods

Population coverage

Usual residents (i.e. excluding visitors) of private dwellings in urban and rural areas of Australia, covering about 98 per cent of the people living in Australia. Private dwellings are houses, flats, home units, caravans, garages, tents and other structures that are used as places of residence at the time of interview. Long-stay caravan parks are also included. These are distinct from non-private dwellings which include hotels, boarding schools, boarding houses and institutions. Residents of non-private dwellings are excluded. The survey also excludes:

- households which contain members of non-Australian defence forces stationed in Australia;
- households which contain diplomatic personnel of overseas governments;
- households in remote and sparsely settled areas of the Northern Territory, accounting for about 20% of the population in the Northern Territory.

Sample size

In each month in 2000–01 a sample of approximately 650 dwellings was selected for the SIHC from the responding households in the MPS. Over the year, this resulted in

approximately 15,500 persons over the age of 15 being included in the sample and, of these, about 85% responded. The final sample consists of 6,786 households, comprising 13,193 persons 15 years old and over.

Sampling design

The sample for the income survey is a sub-sample of private dwellings included in the ABS Monthly Population Survey (MPS). The MPS sample is a multistage selection of private dwellings and a list sample of other dwellings. Each selected dwelling is included in the MPS for eight consecutive months. This is achieved by dividing the sample into eight separate groups and rotating out one group each month. In any particular month, the sample selected for the SIHC is one-sixth of the last rotation group within the MPS sample.

C. Data collection and acquisition

Experienced ABS interviewers collect income information through a personal interview of each person aged 15 or over in the selected dwelling. Housing costs data are collected from all single persons over 15 and the first partner to be interviewed in a married or de facto couple. A separate questionnaire is used for each person.

The information is generally collected after all of the LFS data are obtained. Members of the household not present at the first contact are interviewed later, by telephone if possible. For a telephone call-back interview, the interviewer leaves behind some explanatory material about the survey and a set of prompt cards to help the respondent answer those questions which require him or her to indicate one or more answers from a prepared list of possible choices.

Persons with income from their own business who do not know their income are asked if the ABS can contact them at a later date when records might become available, or, alternatively, written permission is obtained for the ABS to contact the accountant or other financial manager directly to obtain the missing information.

Information about personal and household characteristics, including the labour force status, is collected as part of the LFS, and then combined with the income and housing info during office processing of the survey.

The 2000-01 survey was conducted from July 2000 to June 2001.

D. Definition of the survey units

Income is collected for each individual aged 15 years or more, and individuals are then aggregated in income units and households. Data are presented for individuals, income

units and households. While the income unit was the unit of analysis used in the 1994–95 to 1999–2000 issues of survey results, starting from the 2000-01 issue, the person is used as the unit of analysis with persons mostly described according to the characteristics of the household to which they belong. The income variable most relevant to an individual's economic well-being has thus shifted from income unit income to household income.

An *income unit* is defined as one person or a group of related persons within a household, whose command over income is assumed to be shared. Income sharing is assumed to take place within married (registered or de facto) couples, and between parents and dependent children.

A *family* is defined as two or more people, one of whom is at least 15 years of age, who are related by blood, marriage (registered or de facto), adoption, step or fostering, and who usually live in the same household. A separate family is formed for each married couple, or for each set of parent-child relationships where only one parent is present. A family can contain several income units, as non-dependent children are treated as separate income units.

A *household* is defined as a group of related or unrelated people who usually live in the same dwelling and make common provision for food and other essentials of living; or a lone person who makes provision for his or her own food and other essentials of living without combining with any other person. Lodgers who receive accommodation only (not meals) are treated as a separate household. Boarders who receive accommodation and meals, are treated as part of the household.

The *reference person* for each household is chosen by applying, to all household members aged 15 and over, the selection criteria below, in the order listed, until a single appropriate reference person is identified:

- the person with the highest tenure when ranked as follows: owner without a mortgage, owner with a mortgage, renter, other tenure;
- one of the partners in a registered or de facto marriage, with dependent children;
- one of the partners in a registered or de facto marriage, without dependent children;
- a lone parent with dependent children
- the person with the highest income;
- the eldest person.

For example, in a household containing a lone parent with a non-dependent child, the person with the highest tenure will become the reference person. If the non-dependent child is an owner with a mortgage and the lone parent lives in the dwelling rent free, the non-dependent child will become the reference person. If both individuals have the same tenure, the one with the higher income will become the reference person. However, if both individuals have the same income, the elder will become the reference person.

E. Contents

Incomes

Income refers to regular and recurring cash receipts from employment, investments and transfers from government, private institutions and other households. Gross income is the sum of the income from all these sources before income tax and the Medicare levy have been deducted. This differs from the household income definition used in the Australian System of National Accounts (ASNA) (see appendix to the 1997–98 issue of "Household Income and Income distribution"). Comparison of 2000–01 SIHC and ASNA data indicates that the relationship between the two estimates has not changed significantly since 1997–98.

Sources from which income may be received include:
- wages and salaries (whether from an employer or own corporate enterprise);
- profit/loss from own unincorporated business (including partnerships);
- investment income (interest, rent, dividends, royalties);
- government cash transfers (pensions, allowances, benefits);
- private cash transfers (e.g. superannuation, regular workers' compensation, income from annuities and child support).

Receipts which are excluded from income because they are not regular or recurring cash payments include the following:
- income in kind including employee benefits such as the provision of a house or a car;
- employer contributions to pension and superannuation funds;
- capital transfers such as inheritances and legacies, maturity payments on life insurance policies, lump sum compensation for injuries or other damage;
- capital gains and losses.

The aged persons' savings bonus and self-funded retirees' supplementary bonus, paid as part of the introduction of The New Tax System in 2000–01, are regarded as capital transfers as they were designed to help retired people maintain the value of their savings and investments following the introduction of the GST. However, the one-off payment to seniors announced in the May 2001 Budget and paid in 2000–01 is included as income as it was primarily a supplement to existing income support payments.

Reference period – Income in SIHC is collected for two reference periods, the financial year preceding the date of interview and a much shorter one centered around interview time, i.e. at, or close to, the time of interview (current income). The survey is conducted over twelve months of each year. Estimates of current weekly income derived from the survey do not refer to any given week, but are representative of weekly income during the survey year.

Gross and net income - Income as collected in the survey is gross income before the deduction of personal income tax and Medicare levy. Data on income tax and Medicare levy paid are not collected in the survey. Instead, the ABS models the amount of income

tax (including the Medicare levy) payable by individuals according to the relevant taxation criteria, and using the income and characteristics of individuals and family members as reported in the survey. While the survey data are not sufficiently comprehensive to enable the calculation of exact amounts of tax payable, the model does provide good proxy estimates.

Housing information

Information about housing collected in the survey includes housing tenure; housing costs; and characteristics of the dwelling, such as estimated sale value, value of mortgages and secured loans outstanding, type of dwelling structure and number of bedrooms.

Socio-demographic and labour force information

As the respondents to the SIHC have already completed the LFS, information will have been collected on the household and family characteristics or residents within the dwelling. These characteristics include standard demographic items such as age and sex, marital status, birthplace and year of arrival for those born outside Australia. It includes the standard labour force items such as labour force status, occupation and industry of employer. In addition the SIHC collects information on the highest educational qualification and the current study status from each person and the usual number of hours worked in the first and second jobs.

F. Quality of data

Response rate

In each month in 2000–01 a sample of approximately 650 dwellings was selected for the SIHC from the responding households in the MPS. Over the year, this resulted in approximately 15,500 persons over the age of 15 being included in the sample and, of these, about 85% responded. The final sample consists of 6,786 households, comprising 13,193 persons 15 years old and over.

Data processing

A combination of clerical and computer-based systems is used to process the data. Clerical edits are initially applied by interviewers to ensure completeness and consistency of the questionnaires. After clerical checks and coding are complete, information, excluding name and address, from the questionnaire is entered into the computer. Information collected as part of the LFS is merged from the relevant computer files. An extensive range of computer edits is applied to each record to check that logical sequences have been followed; nothing has been incorrectly included or excluded on the questionnaire; all the necessary items are present; and that specific values lie within valid

ranges. The edits are designed to detect errors which could have occurred during the interview, coding and data entry. Amendments are made to records on the computer file as required. The information is stored on a computer output file in the form of data items. The data items can be formed from answers to single questions, or derived from answers to several questions.

Imputation

Fully non-responding households are those selected for the survey but from which no information is included in the survey results. They include:
- those affected by death or illness of a household member;
- those in which more than half of the persons over 15 in the household did not respond because they could not be contacted, had language problems or refused to participate.

Partial response occurs when:
- some items of data in a schedule are missing because a person is unable or unwilling to provide the data;
- for a household, not every person over 15 residing in the household responds but at least half of these persons provide data.

In the first case of partial response above, the data provided are retained and the missing data are imputed by replacing each missing value with a value reported by another person (referred to as the donor). For the second type of partial response, the data for the persons who did respond are retained, and data for each missing person are provided by imputing data values equivalent to those of a fully responding person (donor). Imputation using donor records is also applied for fully non-responding households that comprise one person or a sole parent whose children are all under the age of 15. Information about the household composition is obtained from the MPS.

Donor records are selected by matching information on sex, age and labour force characteristics of the person with missing information. As far as possible, the imputed information is an appropriate proxy for the information that is missing. Depending on which values are to be imputed, donors are chosen from the pool of individual records with complete information for the block of questions where the missing information occurs.

Of the 6,786 households in the final sample, income information was imputed for 243 households comprising one adult or one adult with children under 15 years old, and was imputed for one or more persons in 201 partially responding multi-person households.

Imputation of one-off payments to seniors - Certain cash receipts that should be regarded as income are not reported in SIHC due to their irregular or one-off nature. For example, annual wage or salary bonuses will not be reported by householders as part of their 'usual' cash income. While these types of income are routinely excluded from SIHC income measures, their exclusion is unlikely to affect comparisons over time unless the scale and distribution of such payments to householders changes. However, as noted in paragraph

18 above, in 2000–01 about $600m was paid in government cash transfers as part of the one-off payment to seniors to supplement the age pension or Department of Veterans' Affairs service pension and therefore should be treated as income. The payment to eligible individuals was $300 each, representing 3% of a single person's full age pension in 2000–01, and this amount has been added to the income of all respondents who were of age pension age and who reported receiving any government income support payment.

Weighting procedure

Expansion factors, or weights, are values by which information for the sample is multiplied to produce estimates for the whole population. From this survey, estimates are produced referring to persons, to income units (although these are not included in this publication) and to households, and the weights are calculated so that each person in an income unit or household has the same weight and that weight is also used for the income unit and household.

Final weights are calculated through an iterative procedure in which initial weights are adjusted by a calibration process to ensure that survey estimates conform to independently estimated benchmarks. The initial weights are equal to the inverse of the probability of selection in the survey, with initial person weights being equal to initial household weights.

Four types of benchmarks are used in the calibration of the final weights:

- numbers of persons aged 15 and over;
- numbers of children under age 15;
- numbers of households;
- for 1999–2000 and 2000–01 estimates, the value of government benefit cash transfers.

Person benchmarks for persons aged 15 and over are estimates of the number of people in each state and territory by age and sex, the number of people in each state and the ACT by labour force status and the number of people in each state living in the capital city or the balance of the state.

A separate set *of benchmarks is used for children* under 15, since there are not individual person records for them in the survey. Information about children is recorded on household records, however, and benchmarks for the number of children aged 0–4 and aged 5–14 are used for each state and territory.

Numbers of households are calibrated to benchmarks for total Australia with respect to household composition (based on the number of adults (1, 2 or 3+) and whether or not the household contains children).

The person and household benchmarks are based on estimates of numbers of persons and households in Australia. The benchmarks are adjusted to include persons and households

residing in private dwellings only and therefore do not, and are not intended to, match estimates of the Australian resident population published in other ABS publications.

The fourth type of benchmark relates to *income from social security transfers*, and is only used for 1999–2000 and 2000–01. The benchmark was introduced for those years because, without it, the survey estimates of income from government benefit cash transfers account for a declining proportion of aggregate social security payments reported by the Department of Family and Community Services and the Department of Veterans' Affairs. Extensive investigations could not identify any specific reasons for the decline, indicating that it is likely to be associated with differences between the characteristics of people who respond to the survey and the characteristics of those who do not respond. This type of problem is sometimes called non-response bias, and introducing additional benchmarks is a means of addressing it. The benchmark introduced in this case ensured that the survey estimate of government benefit cash transfers is maintained at a proportion of aggregate benefit cash transfers that is consistent with the proportion achieved between 1994–95 and 1997–98.

Reliability of estimates

Two types of error are possible in an estimate based on a sample survey: non-sampling error and sampling error.

Non-sampling error

Non-sampling error refers to inaccuracies that may occur because of imperfect reporting by respondents, bias resulting from an inability to obtain data from all households, reporting errors and errors made in processing the data. They can occur whether the estimates are derived from a sample or from a complete collection. It is not possible to quantify non-sampling error, but every effort is made to reduce it to a minimum. This is done by careful design of questionnaires, intensive training and supervision of interviewers, asking respondents to refer to records whenever possible and by extensive editing and quality control checking at all stages of processing. In editing, particular attention was paid to verification of extreme values reported by respondents and to values outside the range considered plausible. Where possible, records were checked to ascertain their accuracy and amended using other information provided by respondents. However, it was not possible to amend all the records or all the data items which may have suspect data. Where there was no additional information available, data items were not changed.

Sampling error

Sampling error arises because the estimates are based on a sample of possible observations and so will differ from estimates that would have been produced if all households had been included in the survey. One measure of the likely difference is given by the standard error (SE), which indicates the extent to which an estimate might have varied because only a sample of dwellings was included. There are about two chances in three (67%) that a sample estimate will differ by less than one SE from the number that

would have been obtained if all dwellings had been included, and about 19 chances in 20 (95%) that the difference will be less than 2 SEs. Another measure of the likely difference is the relative standard error (RSE), which is obtained by expressing the SE as a percentage of the estimate.

For estimates of population sizes, the size of the SE generally increases with the level of the estimate, so that the larger the estimate the larger the SE. However, the larger the sampling estimate the smaller the SE in percentage terms (RSE). Thus, larger sample estimates will be relatively more reliable than smaller estimates. Only estimates with RSEs of 25% or less are considered reliable for most purposes. Estimates with RSEs greater than 25% but less than or equal to 50% should be used with caution, while estimates with RSEs greater than 50% are considered too unreliable for general use and should only be used to aggregate with other estimates to provide derived estimates with RSEs of 25% or less.

Each record on the Confidentialised Unit Record file (CURF) contains 30 'replicate weights' in addition to the 'main weight'. The purpose of these replicate weights is to enable calculation of the RSE on each estimate produced. The basic idea behind the replication approach is to select subsamples repeatedly (30 times) from the whole sample. For each of these subsamples the statistic of interest is calculated. The variance of the full sample statistic is then estimated using the variability among the replicate statistics calculated from the subsamples. As well as enabling variances of estimates to be calculated relatively simply, replicate weights also enable unit record analyses such as chi-square tests and logistic regression to be conducted which take into account the complex sample design.

There are various ways of creating replicate subsamples from the full sample. The replicate weights produced for the 2000–01 SIHC have been created using a group jack-knife method of replication. The formulae for calculating the standard error (SE) and relative standard error (RSE) of an estimate using this method are:

$$SE(y) = \text{sqrt} \left((29/30) \Sigma g \, (y(g) - y)2 \right)$$

where

G = 1,..,30 (the no. of replicate groups)

y(g) = weighted estimate, having applied the weights for replicate group g

y = weighted estimate from the full sample.

$$RSE(y) = SE(y)/y * 100\%.$$

It is not clear that the jackknife method will provide good estimates for the variance of quantile boundaries such as the median (see Rao, J.N.K, Wu, C.F.J., and Yue, K (1992) Some recent work on resampling methods for complex surveys, Survey Methodology, Vol 18, pp.209–217). An indirect approach (known as the Woodruff method) is available for estimating the variance of a quantile based on replicate weights (see Sarndal, Swenson, and Wretman: Model Assisted Survey Sampling, Springer-Verlag, 1992).

To enable CURF users to check that they are using the replicate weights correctly, RSEs for estimates other than medians in table 4.1 have been calculated using the group

jackknife method and are included as table 4.2. The RSEs shown for the medians have been calculated using the Woodruff method.

G. Uses of the survey

Publications

Government Benefits, Taxes and Household Income, Australia, 1998–99, cat. no. 6537.0.

Household Expenditure Survey, Australia: User Guide, 1998–99, cat. no. 6527.0, available free of charge from the ABS web site.

Household Expenditure Survey, Australia: Summary of Results, 1998–99, cat. no. 6530.0.

Household Expenditure Survey, Australia: Detailed Expenditure Items, 1998–99, cat. no. 6535.0.

Housing Occupancy and Costs, Australia, 1997–98, cat. no. 4130.0.

Labour Force, Australia, cat. no. 6203.0—issued monthly.

Survey of Income and Housing Costs and Amenities: Income Units, Australia, 1990, cat. no. 6523.0.

Survey of Income and Housing Costs, Australia: User Guide, 1997, cat. no. 6553.0.

Average Weekly Earnings, Australia—Preliminary, cat. no. 6301.0—issued quarterly.

Measuring Wellbeing: Frameworks for Australian Social Statistics, 2001, cat. no. 4160.0.

Measuring Australia's Progress, 2002, cat. no. 1370.0.

Taxation Statistics 2000–01, A summary of taxation, superannuation and child support statistics (Australian Taxation Office).

Occasional Paper No. 1: Income support and related statistics: a 10-year compendium, 1989–1999 (Department of Family and Community Services)

Poverty and Income Distribution

Selected income distribution indicators, equivalised disposable household income

		Current income basis			*Previous financial year basis*			*Difference in % change*
		1994–95	*1999–2000*	*% change*	*1994–95*	*1999–2000*	*% change*	
Mean income per week, in 2000–01 dollars								
Low income[(a)]	$	227	241	6.4	230	249	8.2	1.8
High income[(b)]	$	792	879	10.9	807	917	13.7	2.8
Income shares								
Low income[(a)]	%	10.8	10.5	–2.3	10.7	10.5	–2.5	–0.2
High income[(b)]	%	37.8	38.4	1.6	37.7	38.6	2.5	0.8

Percentile ratios								
P90/P10	ratio	3.77	3.89	3.1	3.9	4.06	4.3	1.2
P80/P20	ratio	2.56	2.64	3.4	2.63	2.64	0.6	–2.7
Gini coefficient	no.	0.302	0.31	2.8	0.302	0.313	3.6	0.8

Source: Household Income and Income Distribution, Australia, 2000-01, ABS, 2003.

(a) Persons in the 2nd and 3rd income deciles after being ranked by their equivalised disposable household income.
(b) Persons in the top income quintile (9th and 10th deciles) after being ranked by their equivalised disposable household income.