6224.0.55.002

Australian Bureau of Statistics

Information Paper

Improvements to Family Estimates from the Labour Force Survey

Australia

2008

Information Paper

Improvements to Family Estimates from the Labour Force Survey

Australia

2008

Brian Pink
Australian Statistician

AUSTRALIAN BUREAU OF STATISTICS

EMBARGO: 11.30AM (CANBERRA TIME) WED 17 SEP 2008

ABS Catalogue No. 6224.0.55.002

Produced by the Australian Bureau of Statistics

INQUIRIES

- For further information about these and related statistics, contact the National Information and Referral Service on 1300 135 070 or Carmel O'Regan on Canberra (02) 6252 6127.

CONTENTS

INTRODUCTION

THE LABOUR FORCE SURVEY

One of the most important Australian Bureau of Statistics (ABS) labour collections is the Labour Force Survey (LFS). The ABS has conducted the LFS since 1960, first as a quarterly collection and then monthly from February 1978. The LFS collects information from the occupants of a sample of approximately 24,000 dwellings to estimate the labour market activity of Australia's resident civilian population aged 15 years and over. The LFS is designed primarily to provide estimates of key labour force statistics for the whole of Australia and, secondarily, for each state and territory.

The LFS statistics of most interest are the monthly estimates of the number of people employed and unemployed, the unemployment rate and the labour force participation rate. Other data collected monthly or quarterly include hours worked, industry and occupation of employed persons, and duration of unemployment.

FAMILY ESTIMATES FROM THE LFS

Since the 1970s, family estimates have been produced as a by-product of the LFS. Family type data are derived from questions which establish the relationships between members of the households selected in the survey.

In the LFS, a family is defined as:

> Two or more persons, 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 are usually resident in the same household. The basis of a family is formed by identifying the presence of a couple relationship, lone parent–child relationship or other blood relationship. Some households will, therefore, contain more than one family.

A 'couple relationship' includes same-sex couples.

The main source of published LFS family estimates is a series of datacubes released annually in *Labour Force Status and Other Characteristics of Families - Electronic Delivery* (cat. no. 6224.0.55.001). Relationship data at the person level are published each month in datacubes in *Labour Force, Australia, Detailed - Electronic Delivery* (cat. no. 6291.0.55.001). Family and relationship tables are also included in the quarterly publication *Australian Labour Market Statistics* (cat. no. 6105.0).

Labour Force Status and Other Characteristics of Families - Electronic Delivery (cat. no. 6224.0.55.001) contains data for each June, on the number of each family type (for example, couple family/lone parent family, with/without dependants), number of dependent children, age group of dependent children and selected labour force information for the family members.

OBJECTIVES OF THIS INFORMATION PAPER

This paper provides information about improvements to be made from October 2008 to family estimates from the LFS and resulting changes to LFS families products. The estimates will make greater use of available information and will be produced using an improved estimation methodology.

Specifically, the paper details:

- the current estimation method;
- the improved estimation method;
- the impact on family estimates resulting from the changed methodology; and
- provides comparisons with other ABS family estimates.

OBJECTIVES OF THIS INFORMATION PAPER *continued*

Finally, a range of graphs are presented which highlights the outcomes from the adoption of the new methodology in terms of family estimates from the LFS.

CURRENT ESTIMATION METHOD

LFS PERSON-LEVEL PROCESSING

Currently, LFS family estimates are a by-product of LFS person-level estimates. To understand the current method of producing family estimates from the LFS, it is first necessary to understand the method of producing person-level estimates from the LFS.

Scope exclusions

The LFS is designed primarily to measure key characteristics of the civilian labour force in Australia. As a result, certain population groups are excluded from LFS person-level estimates, because they do not fall within the definition of the civilian labour force in Australia. Examples are permanent members of the Australian defence forces, children aged less than 15 years, and overseas diplomats living in Australia. People in these population groups are referred to as 'out on scope'.

Selection rule exclusions

People who fail to meet LFS selection rules are excluded from LFS person-level estimates. Selection rules are designed to associate each person in Australia with one and only one dwelling, to ensure each person has only one chance of selection in the survey. For example, persons away from their usual residence for more than six weeks are excluded from the LFS at their usual residence, but if the dwelling they are visiting is selected in the survey, they are included in the survey at that dwelling.

Exclusions due to partial non-response

People for whom key LFS information is missing are excluded from LFS person-level estimates.

Person weights

As part of LFS person-level processing, each person record which is in scope of the survey, fully responding, and meets LFS selection rules is assigned an expansion factor or 'weight' so that the distribution of the weighted sample reflects the in-scope Australian population in each state and territory and at a regional level. The weights are calculated so that they sum to independent estimates of the civilian population aged 15 years and over (referred to as population benchmarks).

Weighting the sample data to the population benchmarks compensates for people who did not fully respond to the survey. However, the population benchmarks explicitly exclude people who are outside the scope of the LFS.

FAMILY WEIGHTING

To produce family estimates, each family is assigned a family weight, so that the responses from the families in the survey sample can be inflated to represent all families in Australia.

Under the current method, the value of the family weight is the harmonic mean of the person weights of all family members aged 15 years and over.

Estimates for each characteristic of interest are then obtained by summing the weights of the families in the sample with that characteristic.

ISSUES WITH THE CURRENT METHOD

There are two main issues with the current method:

- Some families in the LFS sample are not assigned a weight, and therefore make no contribution to LFS family estimates.
- There is no use of independent population benchmarks, such as the Estimated Resident Population (ERP), in assigning the family weight. Therefore, no compensation is made for those families which make no contribution to LFS family estimates.

ISSUES WITH THE CURRENT METHOD *continued*

Further information about these excluded families is provided below.

Families in incomplete households excluded

Under the current method, a family weight is only assigned to families in households where all usual residents of the household aged 15 years and over have contributed to the person-level estimates. Families in households in which one or more members aged 15 years and over have not contributed to the person-level estimates (for example because they are permanent members of the Australian defence forces, did not fully respond to the survey, or failed to meet selection rules) are not assigned a family weight and therefore make no contribution to the estimates. Without the use of independent population benchmarks, no compensation is made for these excluded families.

Families in discrete Indigenous communities excluded

In discrete Indigenous communities, the LFS questionnaire does not collect information about relationships within households, due to the complexity and cost of doing so. Without information about relationships within households, family data cannot be derived from households in discrete Indigenous communities selected in the LFS.

Under the current family estimation method, families in discrete Indigenous communities make no contribution to the estimates, and without the use of independent population benchmarks, no compensation is made for these excluded families.

Effect of excluded families

The combined effect of these exclusions is that the current LFS family estimates do not accurately reflect the Australian population. The exclusion of some families in the LFS sample prior to taking the harmonic mean, without the compensating effects of independent population benchmarks, results in family estimates being lower than they would be if all families were included. Also, the family characteristics of the excluded groups, such as households containing defence force personnel, may be different from the included groups.

The lack of independent benchmarks also means that LFS family estimates are not comparable with other ABS estimates of families. The total number of children, adults and households are not constrained to equal the independent benchmarks, which means LFS family estimates differ from other ABS surveys.

SUMMARY

The current method of producing LFS family estimates excludes:

- families in households in which one or more members aged 15 years and over have not contributed to the person-level estimates (for example because they are permanent members of the Australian defence forces, did not fully respond to the survey, or failed to meet selection rules); and
- families in discrete Indigenous communities.

The exclusion of these families, without the compensating effects of independent population benchmarks, results in family estimates being lower than they would be if all families were included.

IMPROVED ESTIMATION METHOD

IMPROVEMENTS

From October 2008, the method of producing family estimates from the LFS will be improved in the following ways:

- The conceptual scope of the family estimates will widen to include households containing permanent members of the Australian defence forces.
- A wider range of families in the LFS sample will contribute to the family estimates.
- An improved weighting method will be employed, which makes use of independent population benchmarks (of persons and households) to ensure the estimates more closely reflect the Australian population.

SCOPE OF THE FAMILY ESTIMATES

Australian defence force personnel to be included

The current method of producing family estimates begins with those records used in person-level processing. The aim of person-level processing is to produce employment estimates of the civilian population of Australia. Permanent members of the Australian defence forces are explicitly excluded from the scope of LFS person-level estimates. Because LFS family estimates were derived from the person-level estimates, the families of Australian defence force personnel have until now been excluded from the scope of LFS family estimates.

Under the improved method, the scope of the improved family estimates is independent of person-level processing. The improved method will provide estimates of all Australian families who usually live in private dwellings, including the families of Australian defence force personnel.

Non-private dwellings and other continuing exclusions

Persons in non-private dwellings (for example hotels, hospitals and camping grounds) are in scope of LFS person-level estimates, but will continue to be excluded from the scope of LFS family estimates. The LFS does not collect relationship information in non-private dwellings. Most people selected in non-private dwellings in the LFS are staying there temporarily, and usually reside in a private dwelling.

A number of other groups excluded from ERP will continue to be excluded from the scope of LFS family estimates. A household will be excluded from family estimates if any of the usual residents aged 15 years and over are:

- Non-Australian diplomats, non-Australian diplomatic staff or non-Australian members of their household;
- Short term overseas visitors, whose usual residence is outside of Australia and who are staying in Australia for less than 12 months; or
- Members of non-Australian defence forces stationed in Australia or their dependants.

WIDER RANGE OF FAMILIES INCLUDED

Under the improved method, data from all private dwellings will contribute to family estimates, as long as all demographic and family type variables are available. This will provide a more complete picture of family characteristics from the LFS than has been available in the past.

Since the introduction of Computer Assisted Interviewing (CAI) in the LFS in August 2004, basic demographic information has been captured for all persons in selected dwellings including those out on scope of the LFS. Since July 2006, additional information has been collected (for example, whether persons aged 15-24 were in full-time education) for all usual residents of private dwellings, including those out on

WIDER RANGE OF FAMILIES INCLUDED *continued*

scope. These changes allow all households in private dwellings to be coded to family type, and used in the production of family estimates.

With the introduction of the improved method in October 2008, information from the following groups will contribute for the first time to LFS family estimates:

- Households containing permanent members of the Australian defence forces;
- Households containing usual residents of private dwellings who fail to meet normal LFS selection rules (for example, because they are away from their usual residence for more than six weeks); and
- Households which did not fully respond to the survey (as long as all demographic and family type variables are available).

Discrete Indigenous communities

Data on relationships within households in discrete Indigenous communities are not available from the LFS, due to the complexity and cost of collecting this information. Therefore, relationship data from discrete Indigenous communities is not available for use in the production of LFS family estimates.

However, the scope of LFS family estimates is all private dwellings, including private dwellings in discrete Indigenous communities. The use of independent population benchmarks in weighting (explained below) compensates somewhat for the lack of household relationship information in discrete Indigenous communities, by ensuring that the family estimates are consistent with benchmark counts of persons and households for a given area (e.g. Northern Territory).

Any benchmarking process assumes that the unobserved data are represented by the observed data. The improved LFS family methodology assumes that families in discrete Indigenous communities have similar characteristics to all other families in the same geographic area (see Improved Weighting Method below for the geographic areas used in benchmarking). This assumption should be kept in mind, particularly when interpreting data for the Northern Territory, where approximately 20% of the population lives in discrete Indigenous communities. If families in discrete Indigenous communities have different characteristics to families in the rest of the Northern Territory, for example if families are larger, then the number of families in the Northern Territory will be overstated. The improved LFS estimate of families in the Northern Territory are likely to be much closer to the true number of families than the current LFS estimate (see the comparison with other ABS sources of families data in the next section of this publication), but some characteristics of those families may be less reliable.

IMPROVED WEIGHTING METHOD

The improved weighting method to be introduced from October 2008 makes use of independent population estimates, known as benchmarks. The method makes use of benchmarks of both persons and households. The use of benchmarks has three benefits:

- It ensures the estimates are more representative of the Australian population;
- The estimates are more comparable with estimates from other ABS collections; and
- Sampling error is considerably reduced.

Person benchmarks

The estimates of people used in the family estimates will be weighted to match to independent counts of persons for each cross-classification of the following class variables:

Person benchmarks continued

- geographic area - State capital city (e.g. Sydney), balance of State (e.g. New South Wales, excluding Sydney), Northern Territory and Australian Capital Territory (14 classes);
- sex (2 classes); and
- age group - 5 year age groups between 0 and 74, 75 years and over (16 classes).

The combination of these three variables gives 448 benchmark classes (14 x 2 x 16 = 448). The weighted total of persons within each benchmark class is constrained to match the population count in that class. For example, the weights of 15-19 year old males who live in Sydney will be adjusted so that they total the estimated resident population of 15-19 year old males in Sydney.

Household benchmarks

The number of households determined from the processing of family estimates will be weighted to match the independent counts of households in Australia for each cross-classification of the following class variables:

- geographic area - State capital city (e.g. Sydney), balance of State (e.g. New South Wales, excluding Sydney), Northern Territory and Australian Capital Territory (14 classes); and
- household composition - the number (1, 2 or 3+) of usual residents who are aged 15 years or over, and the number (0, 1 or 2+) of usual residents who are aged under 15 years (9 classes).

The combination of these two variables gives 126 benchmark classes (14 x 9 = 126). The weighted total of households within a benchmark class is constrained to match the total number of households in that class.

For example, the weights of households in Sydney with one usual resident aged 15 years or over and one usual resident aged under 15 years will be adjusted so that they total the estimated number of households in Sydney with the same household composition.

The current household benchmarks used in the derivation of family estimates are based on the 2001 Census of Population and Housing. The incorporation of information from the 2006 Census of Population and Housing and a review of the methodology for producing household benchmarks is likely to result in revisions to the household benchmarks and in turn the family estimates.

Computing the family weight

Each family within the scope of family estimates will be assigned an expansion factor or weight, using a form of 'Generalised Regression' (GREG). The outputs of this process are household weights, which are set to the family weights and person weights for the families and individuals in the household. The sum of the final weights within each benchmark class are constrained to equal the benchmark count at both the person level and household level.

Estimates for each characteristic of interest are then obtained by summing the weights of families in the sample with that characteristic.

SUMMARY

To reiterate, the improved estimation method will, through the inclusion of a wider range of families and the use of independent population benchmarks, result in family estimates that more closely reflect the Australian population.

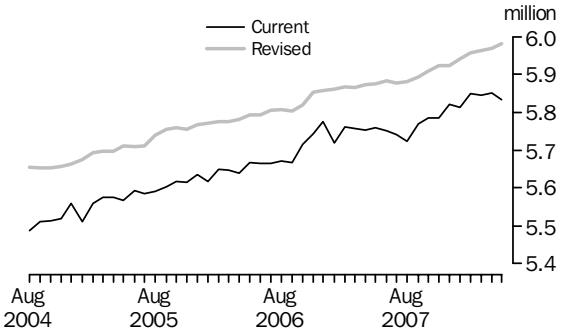
SUMMARY *continued*

The main limitation of the improved estimation method is the necessary assumption that families in discrete Indigenous communities have family characteristics similar to other families in the same area. This may result in some overstatement of families in the estimates.

INCREASED NUMBER OF FAMILIES

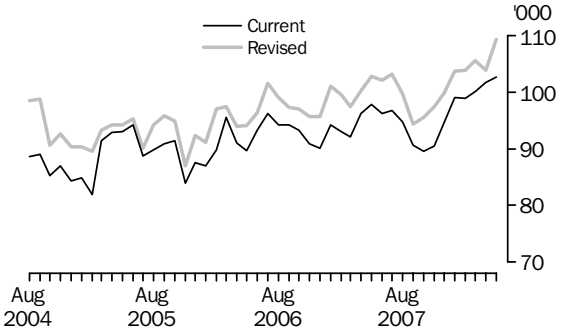
The current estimation method underestimates the total number of families in Australia, because some population groups are excluded from the estimates, and because no independent population estimates (benchmarks) are used. Under the improved estimation method, a larger set of people in the LFS will contribute to the family estimates, and the estimation method will make use of benchmarks for persons and households. As a result, the estimate of the total number of families in Australia will increase on average by 2%. The graph below shows the total number of families in Australia over the period August 2004 to May 2008 under the current estimation method and under the improved estimation method.

FAMILIES, AUSTRALIA



The estimate of the number of families of each type (for example, couple families with dependent children under 15 years of age) will likewise increase, although the size of the increase varies according to family type. The largest proportional increase is for other families (that is, families which are not couple families or lone parent families), which increases on average by 5% (see graph below) albeit from a small base.

OTHER FAMILIES, AUSTRALIA



The graphs at the end of this publication show the estimated number of families of each type for the period August 2004 to May 2008 produced by the current estimation method compared with the improved estimation method.

STATE/TERRITORY LEVEL

The improved estimation method results in a large increase in the total number of families in the Northern Territory, and a small increase in the states and the Australian Capital Territory. The graphs at the end of this publication show the estimated number of families in each state/territory for the period August 2004 to May 2008 produced by the current estimation method compared with the improved estimation method.

STATE/TERRITORY LEVEL *continued*

In the Northern Territory, the improved estimation method produces an estimate of total families which is approximately 40% higher than the current estimate. This is primarily due to the number of families in the Northern Territory being considerably underestimated under the current method (see the previous section on the Current Estimation Method for details). The current method underestimates the number of families in the Northern Territory because:

- households in discrete Indigenous communities are excluded;
- households in which one or more members aged 15 years and over are permanent members of the Australian defence forces are excluded;
- these excluded households are more likely to contain families than other households; and
- the current families estimation method does not make use of benchmarks to compensate for these excluded households.

The improved estimation method provides estimates of higher quality than the current method. The improved method includes households in which one or more members aged 15 years and over are permanent members of the Australian defence forces, and makes use of person and household benchmarks, which compensates somewhat for the lack of household relationship information in discrete Indigenous communities.

COMPARISON WITH OTHER ABS SOURCES OF FAMILIES DATA

Survey of Income and Housing

A key source of ABS data on families is the Survey of Income and Housing (SIH), which is conducted every two years. The SIH collects detailed information about income and personal and household characteristics of persons aged 15 years and over resident in private dwellings. The SIH sample of approximately 10,000 dwellings is drawn from all parts of Australia, except very remote areas, which are excluded from the scope of the survey.

Like the improved LFS family method, the SIH makes use of person and household benchmarks, however, those benchmarks exclude very remote areas to match the survey scope.

The table below compares families data from the SIH with families data from the LFS under the current estimation method and the improved estimation method. The LFS data for each month of 2005/06 have been averaged for comparison with SIH.

The scope of the LFS includes very remote areas of Australia, whereas the scope of the SIH excludes them. Therefore, LFS estimateS of the number of families would be expected to be higher than SIH, to account for the wider scope.

Survey of Income and Housing continued

COMPARISON OF IMPROVED LFS FAMILY ESTIMATES WITH SURVEY OF INCOME AND HOUSING—2005/06

State/Territory	Current LFS	Improved LFS	Survey of Income and Housing	Difference between current LFS and SIH	Difference between improved LFS and SIH
	'000	'000	'000	%	%
New South Wales	1 862.0	1 894.0	1 951.2	–4.6	–2.9
Victoria	1 397.0	1 420.0	1 417.9	–1.5	0.2
Queensland	1 111.0	1 151.0	1 116.3	–0.5	3.1
South Australia	435.0	440.0	432.4	0.6	1.8
Western Australia	556.0	571.0	556.1	—	2.7
Tasmania	140.0	142.0	135.4	3.4	4.9
Northern Territory	38.0	55.0	42.0	–9.4	31.1
Australian Capital Territory	87.0	91.0	90.4	–3.7	0.7
Australia	5 627.0	5 765.0	5 741.6	–2.0	0.4

— nil or rounded to zero (including null cells)

The table shows that the current LFS families estimates are generally lower than those from SIH, despite the wider scope. The largest difference is in the Northern Territory, where the current LFS estimate of the number of families is 9.4% lower than the SIH estimate.

The table also shows that the improved LFS families estimates are generally higher than the SIH estimates for all states/territories, except New South Wales, where the improved LFS family estimate is 2.9% lower than SIH. In the Northern Territory, the improved LFS family estimate is 31.1% higher than SIH. This difference is largely attributable to the large proportion of the population in the Northern Territory who live in very remote areas, which are in scope of LFS family estimates, but excluded from the scope of SIH.

SUMMARY

The improved LFS family estimates compare favourably with those from SIH, and have the following advantages:

- Very remote areas of Australia are in scope of LFS family estimates, but are excluded from the scope of SIH.
- LFS family estimates are produced from a sample of approximately 24,000 dwellings, whereas SIH estimates are produced from a sample of approximately 10,000 dwellings. LFS family estimates will therefore have lower sampling error.
- From October 2008, LFS family estimates will be available on a monthly basis, whereas SIH estimates are available every 2 years.

Family Characteristics and Transitions

Another source of ABS data on families is the Family Characteristics and Transitions Survey (FCTS), conducted in 2006-07. The FCTS was part of the Multi-Purpose Household Survey (MPHS), which is collected from a fraction of households who are in their final month of participation in the Monthly Population Survey (the survey vehicle which also includes the LFS).

The FCTS collected information on household and family composition, with a particular focus on families with children aged 0-17 years. The scope of the FCTS was usual

Family Characteristics and Transitions continued

residents of private dwellings in all parts of Australia except very remote areas. The estimates were derived from a sample of approximately 13,000 dwellings.

Like the improved LFS family method, the FCTS makes use of person and household benchmarks, however, those benchmarks exclude very remote areas to match the survey scope.

The table below compares families data from the FCTS with families data from the LFS under the current estimation method and the improved estimation method. LFS data for the 10 months of FCTS enumeration in 2006/07 have been averaged for comparison.

The scope of the LFS includes very remote areas of Australia, whereas the scope of the FCTS excludes them. Therefore, LFS estimates of the number of families would be expected to be higher than FCTS, to account for the wider scope.

COMPARISON OF IMPROVED LFS FAMILY ESTIMATES WITH FAMILY CHARACTERISTICS AND TRANSITIONS—2006/07

State/Territory	Current LFS	Improved LFS	Family Characteristics and Transitions	Difference between current LFS and FCTS	Difference between improved LFS and FCTS
	'000	'000	'000	%	%
New South Wales	1 894.3	1 920.4	1 940.0	–2.4	–1.0
Victoria	1 425.8	1 443.1	1 475.0	–3.3	–2.2
Queensland	1 143.9	1 169.9	1 180.0	–3.1	–0.9
South Australia	436.0	444.8	448.0	–2.7	–0.7
Western Australia	572.3	588.7	584.0	–2.0	0.8
Tasmania	139.8	142.0	142.0	–1.5	—
Northern Territory	40.1	54.9	43.0	–6.8	27.6
Australian Capital Territory	87.8	92.2	94.0	–6.6	–1.9
Australia	5 739.9	5 855.9	5 905.0	–2.8	–0.8

— nil or rounded to zero (including null cells)

The table shows that the current LFS families estimates are generally lower than those from FCTS, despite the wider scope. The largest difference is in the Northern Territory and the Australian Capital Territory, where the current LFS estimate of the number of families is 6.8% and 6.6% lower than the FCTS estimate.

A possible contributing factor to the differences for the Northern Territory and the Australian Capital Territory is the fact that they have a higher proportion of households containing a person outside the scope of the LFS (for example, permanent members of the Australian defence forces) than the states. Since the current LFS family estimation method excludes families with persons out of scope of the LFS and the method makes no compensation through benchmarking, the underestimation of the number of families is higher in the territories.

In the states, the Australian Capital Territory, and at the Australia level, the improved LFS families estimates are similar to the FCTS estimates, despite the LFS having a wider scope. This appears to be related to the higher proportion of family households (as opposed to non-family households such as single-person households and group households) in the FCTS sample compared to the LFS sample under the improved

Family Characteristics and Transitions continued

method. In the FCTS, a family did not contribute to the estimates if one or more members of the household did not contribute to LFS person-level estimates (for example, the person failed to meet LFS selection rules). Non-family households are more likely to have one or more members of the household who did not contribute to the LFS person-level estimates, thus increasing the proportion of family households in the FCTS sample.

In the Northern Territory, the improved LFS family estimate is 27.6% higher than FCTS. This difference is largely attributable to the large proportion of the population in the Northern Territory who live in very remote areas, which are in scope of LFS family estimates, but excluded from the scope of FCTS.

SUMMARY

The improved LFS family estimates compare favourably with those from FCTS, and have the following advantages:

- Very remote areas of Australia are in scope of LFS family estimates, but are excluded from the scope of FCTS.
- LFS family estimates are produced from a sample of approximately 24,000 dwellings, whereas FCTS estimates are produced from a sample of approximately 12,000 dwellings. LFS family estimates therefore have lower sampling error.
- From October 2008, LFS family estimates will be available on a monthly basis, whereas FCTS estimates are expected to be available every 3 years.

Census

Another important source of ABS families data is the Census of Population and Housing, conducted every five years. The main advantages of Census data are the fine level of geographic detail available, the complete coverage of all geographic areas, including very remote areas, and the increased reliability resulting from a complete Census of the Australian population, rather than a sample survey.

While the Census aims to collect data on every person and household in Australia, there are a number of limitations to families data from the Census, including:

- The difficulty of establishing all the relationships which exist in a household, or to identify whether more than one family is living in the dwelling, from the data provided in response to a generalised questionnaire format required by a census.
- A maximum of three families are able to be identified within a single dwelling, in accordance with existing ABS standards. While this may have only a small effect on total family numbers generally, the impact may be more significant among those population groups which are more likely to live in multi-generational households or with larger numbers of extended family members, such as the Indigenous population.
- Family relationships are only identified for the 94.4% of persons who are in their usual residence on Census night. This means that entire families who were away from home on Census Night (on holiday, for example) will not be counted as such in the Census (however, individual family members remain in the person counts as visitors to a dwelling on Census Night).
- For persons imputed into dwellings for which no form was received, there are no relationship data and no families can therefore be identified.

Census continued

- It is not possible to enumerate all households in the Census. Underenumeration in the 2006 Census is estimated to be 2.7%.

The table below compares families data from the Census with families data from the LFS under the current estimation method and the improved estimation method. LFS data are for August 2006 for comparison.

COMPARISON OF IMPROVED LFS FAMILY ESTIMATES WITH 2006 CENSUS OF POPULATION AND HOUSING—August 2006

State/Territory	Current LFS '000	Improved LFS '000	2006 Census of Population and Housing '000	Difference between current LFS and Census %	Difference between improved LFS and Census %
New South Wales	1 894.0	1 913.0	1 716.0	10.4	11.5
Victoria	1 407.0	1 430.0	1 294.0	8.7	10.5
Queensland	1 105.0	1 153.0	1 032.0	7.1	11.7
South Australia	438.0	442.0	407.0	7.6	8.6
Western Australia	558.0	579.0	511.0	9.2	13.3
Tasmania	142.0	143.0	128.0	10.9	11.7
Northern Territory	39.0	55.0	43.0	–9.3	27.9
Australian Capital Territory	89.0	92.0	84.0	6.0	9.5
Australia	5 672.0	5 807.0	5 219.0	8.7	11.3

The table shows that the current LFS families estimates are generally higher than family counts from the Census, except in the Northern Territory, where the current LFS family estimate is 9.3% lower than Census.

Under the improved LFS family estimation method, the difference between the LFS family estimates and Census family counts is slightly greater than under the current method, except in the Northern Territory where the LFS family estimate is 27.9% higher than the Census family count. This is likely to be related to the limitations of Census families data, listed above. Census underenumeration is higher in the Northern Territory (7.6%). Other limitations of Census data may also be more pronounced in the Northern Territory than elsewhere, due to its large Indigenous population.

SUMMARY

The improved LFS estimates of the number of families are considerably higher than the family counts from the Census. Although Census counts provide data at a very fine level of geographic detail, they appear to understate the number of families due to:

- limitations around the general Census questionnaire format;
- existing ABS standard of applying a maximum of three families within a single dwelling;
- families are identified only if they are at their usual residence on Census night; and
- there are no relationship data for persons imputed into dwellings.

These limitations understate the number of families, especially in the Northern Territory.

BREAK IN TIME SERIES

The improved estimation method and consequent increases in the size of the estimates will cause a break in time series. To enable comparisons to be made over time, revised historical data for the period August 2004 to August 2008 will be made available at the same time as data for September 2008 are released (16 October 2008). Data will be released monthly thereafter. Data for months prior to August 2004 cannot be revised using the improved estimation method, as insufficient family information was collected at that time.

REDUCED VOLATILITY

The graphs above, and at the end of this publication, show that the improved family estimates produce a time series which exhibits less volatility than the current time series. This is primarily due to the use of independent population benchmarks.

MONTHLY FREQUENCY

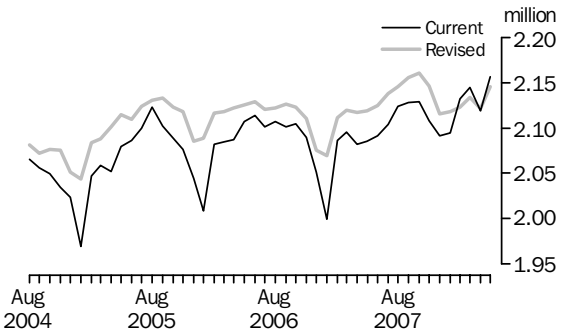
From October 2008, family estimates will be released monthly, instead of annually. The data will be released in the families datacubes (*Labour Force Status and Other Characteristics of Families - Electronic Delivery*, cat. no. 6224.0.55.001). Each of the five datacubes (FA1 to FA5) will be split into two separate datacubes, one for the current methodology and one for the improved methodology. Using FA1 as an example, the FA1_jun94 datacube will contain annual data as currently published from June 1994 to June 2004, and the FA1_aug04 datacube will contain monthly data from August 2004 onwards calculated using the new methodology. These two datacubes will replace the current FA1 datacube, which will be withdrawn.

Seasonal patterns

Care should be taken when comparing families data from month to month. Some series appear to have a strong seasonal pattern and no seasonal adjustment process has been applied to the data. Changes from month to month should be interpreted in the context of year to year change, to ensure that normal monthly fluctuations are interpreted in the context of any seasonal pattern.

An example is couple families with dependants (see graph below). The definition of 'dependant' includes all family members under 15 years of age; family members aged 15-19 attending school or aged 15-24 attending a tertiary educational institution full time (except those classified as husbands, wives or lone parents).

COUPLE FAMILIES WITH DEPENDANTS UNDER 25 YEARS, AUSTRALIA



The graph shows the estimate of the number of couple families with dependants drops considerably in January of each year, which is more noticeable under the current LFS family method than under the improved method. January enumeration of the LFS falls during the holiday period for both school and tertiary students, increasing the likelihood

Seasonal patterns continued

of a family with dependants staying away from their usual residence at the time of LFS enumeration. Families with dependants are therefore less likely than other family types to contribute to family estimates in January. Without the use of benchmarks to compensate for these non-responding families, the underestimation of couple families with dependants under the current LFS family method is especially noticeable in January each year.

The drop in January each year is also noticeable under the improved LFS family method. This may be partly related to the increased likelihood of couple families with dependants to be staying away from their usual residence at the time of LFS enumeration, but also to uncertainty surrounding the full time status of students aged 15-24 years. To be classed as a 'dependant' a person aged 15-24 years must be either at school or attending a tertiary educational institution full time. In December, January and February, a number of people aged 15-24 years are reported as having a full time student status of 'not known/unclear', and a lower proportion are reported as studying full time. This results in a lower estimate of couple families with dependants in these months.

'NOT DETERMINED' LABOUR FORCE STATUS

The new method of producing family estimates uses all LFS households where sufficient family information has been provided. A small proportion of respondents in those households are outside the scope of LFS person-level estimates, for example because they are permanent members of the Australian defence forces, or because they fail to meet normal LFS selection rules. No labour force information is collected for persons out of scope of LFS. Therefore there will be a small proportion of persons included in LFS family estimates for whom no labour force information is available. In the published data, these persons will be assigned to one of two new categories: 'Not determined, defence force personnel' or 'Not determined, other'.

'Not determined, defence force personnel' will denote that a particular characteristic (such as labour force status) was not determined, because the family member was a permanent member of the Australian defence forces, and therefore outside the scope of the LFS. 'Not determined, other' will denote that a particular characteristic could not be determined for other reasons, such as the family member failed to meet selection rules, or did not respond fully to the survey.

Within the families datacubes (*Labour Force Status and Other Characteristics of Families - Electronic Delivery*, cat. no.6224.0.55.001) the 'Not determined' categories will apply to a small proportion of records when classifying by labour force status, status of employment and duration of unemployment.

The 'Not determined' categories will also apply to other items collected which are not currently included in the families datacubes, but are available on request, such as hours worked and occupation.

The 'Not determined' categories give users the flexibility to decide how they are to treat defence force families and other types of previously excluded families.

EFFECT ON SAMPLING ERROR

The estimates produced from the LFS, as from all surveys, are subject to sampling error; that is, they may differ from the estimates that would have been produced if all dwellings had been included in the survey. The most common way of quantifying sampling error is to calculate the standard error for the estimate. The standard error indicates the extent to which an estimate might have varied by chance because only a sample of dwellings was included in the survey.

The improved estimation method is expected to reduce the standard error on family estimates by up to 60%, compared to the current estimation method. The decrease in standard error is mainly attributable to the use of independent population estimates (benchmarks). The size of reduction in the standard error for a given estimate is largely dependent on the correlation between the estimate itself and the independent population estimates. For example, the standard error on estimates of the total number of families for Australia and for each State/Territory will decrease considerably, because the total number of families is highly correlated to the number of persons and households, which are used as benchmarks. There is less benefit in terms of standard error for estimates of each type of family, because the correlation to the benchmarks is less strong.

Independent of the improved estimation method, however, is a reduction in the size of the LFS sample from July 2008. Further information about the sample reduction can be found in the April and May releases of *Labour Force, Australia* (cat. no. 6202.0). From July 2008 the sample size of the LFS was reduced by 24% when compared with the June 2008 sample. The sample reduction increased standard errors on all LFS estimates, including family estimates from July 2008 by approximately 15%, compared to the standard errors which would have been observed without the sample reduction. The relative change in sample size varies across the States and Territories, with corresponding impacts on standard errors.

STANDARD ERROR MODEL

Due to technical limitations, it is not possible to publish the standard error associated with every possible cross-tabulation in the LFS datacubes. Instead, the datacubes annotate those estimates where the standard error is 25% or more of the estimate itself. The standard errors used for the annotations are based on a standard error model, which provides an indication of the standard error for any estimate without the need to access unit record data. The standard error models are produced by statistically modelling the standard errors as a function of the estimate itself. Standard errors derived from a statistical model are an approximation of the standard errors calculated directly from the data.

Currently, the standard error model underlying the families datacubes is the model used for LFS estimates of employed persons. Since the introduction of composite estimation in May 2007 (for details, see the 2007 issue of *Forthcoming Changes to Labour Force Statistics*, cat. no. 6292.0) and other changes, the employed persons model does not give a very accurate indication of the true standard errors associated with the family estimates. From October 2008, the families datacubes will use an improved standard error model, derived specifically for use with family estimates. The improved standard error model will also be made available to users in the spreadsheet *Labour Force Survey Standard Errors, Data Cube* (cat. no. 6298.0.55.001).

STANDARD ERROR MODEL *continued*

The table below shows the point at which the standard error is 25% or more of the estimate itself for June 2008 (before the sample reduction) and July 2008 onwards (after the sample reduction). Estimates which are smaller than these estimates are associated with standard error too high for most practical purposes.

For example, in June 2008, the LFS estimate of the number of same sex couple families with dependants in Australia is 2,700. This figure is below the Australia-level cut-off of 5,831, and therefore associated with standard errors too high for most practical purposes.

25% RELATIVE STANDARD ERROR CUT-OFFS, FAMILY ESTIMATES

	June 2008	*July 2008 onwards*
New South Wales	6 842	8 912
Victoria	6 130	7 944
Queensland	5 657	7 366
South Australia	3 240	4 173
Western Australia	4 387	5 684
Tasmania	1 813	2 327
Northern Territory	1 640	2 047
Australian Capital Territory	1 857	2 381
Australia	5 831	7 704

FEWER RECORDS WITH RELATIONSHIP 'NOT DETERMINED'

Relationship data at the person level are published each month in datacubes in *Labour Force, Australia, Detailed - Electronic Delivery* (cat. no. 6291.0.55.001). Currently the relationship of a usual resident of a household is coded as 'Not determined' if any member of the household is out of the scope of the survey (for example, because he/she is a permanent member of the Australian defence forces), fails to meet selection rules (for example, because he/she is away for more than six weeks) or fails to fully respond to the survey. This 'Not determined' category accounts for approximately 2% of persons in the survey. The improved method of producing family estimates means that most of these previously excluded records can now be included in families processing, and their relationship can now be correctly coded.

The result of this change is that the number of records assigned to the 'Relationship not determined' category will decrease and the number of records assigned to other relationship categories will correspondingly increase. This will cause a very minor break in series.

EFFECT ON PRODUCTS

STANDARD PRODUCTS AFFECTED

The standard products affected by the new families methodology are:

- *Labour Force Status and Other Characteristics of Families - Electronic Delivery* (cat. no. 6224.0.55.001);
- *Labour Force, Australia, Detailed - Electronic Delivery* (cat. no. 6291.0.55.001); and
- *Australian Labour Market Statistics* (cat. no. 6105.0).

From the release on 16 October 2008, *Labour Force Status and Other Characteristics of Families - Electronic Delivery* will consist of two separate sets of datacubes. The first set will contain family estimates for each month from August 2004 onwards produced using the improved method. The second set will contain family estimates for June of previous years produced using the current method.

In the datacubes for August 2004 onwards, the labour force status variable will have the two new 'Not determined' categories ('Not determined - defence force personnel' and 'Not determined - other'). Variables related to labour force status, including status of employment of husband or head of family, will likewise include the two new 'Not determined' categories.

From the release of September 2008 data on 16 October 2008 in *Labour Force, Australia, Detailed - Electronic Delivery* (cat. no. 6291.0.55.001), relationship in household data appearing in datacubes FM1 to FM4 will have fewer 'relationship not determined' responses for usual residents of households, as many of these records will have been coded to their correct category. Data for months prior to September 2008 will have a slightly higher number of 'relationship not determined' responses than data for September 2008 onwards. Data for August 2004 to August 2008 will be revised with the five-yearly revision of LFS data in early 2009, resulting in a consistent time series for August 2004 onwards.

From the January 2009 issue of *Australian Labour Market Statistics* (cat. no. 6105.0), table 1.6 of this publication will reflect the reduced number of 'relationship not determined' responses as these records will have been recoded to their correct category. The labour force status of family members appearing in table 1.7 will incorporate the two new 'Not determined' categories.

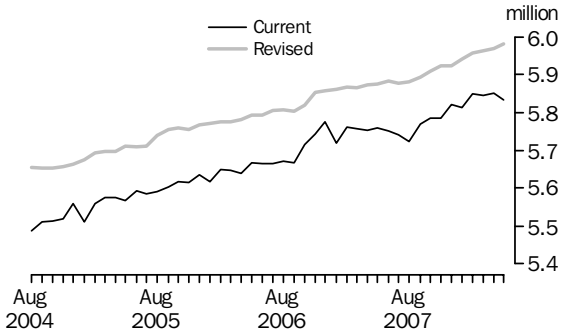
FUTURE DEVELOPMENTS

At this stage, while estimates will be revised based on the new methodology, there will be no change to the format or structure of LFS families products. However, in the near future, ABS intends to review whether the standard families products can be improved, and whether additional families products are required.

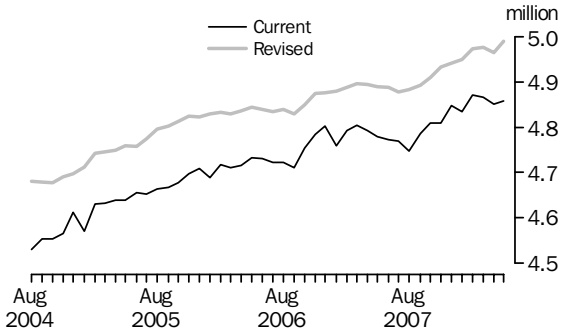
GRAPHS

AUSTRALIA BY FAMILY TYPE

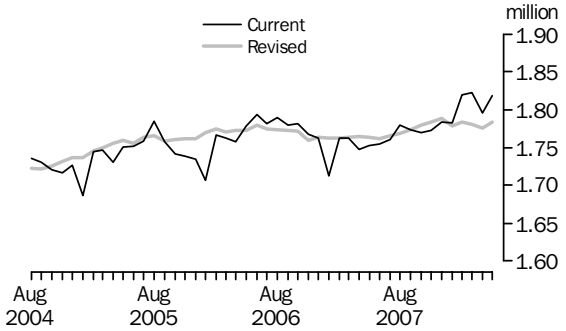
FAMILIES, AUSTRALIA



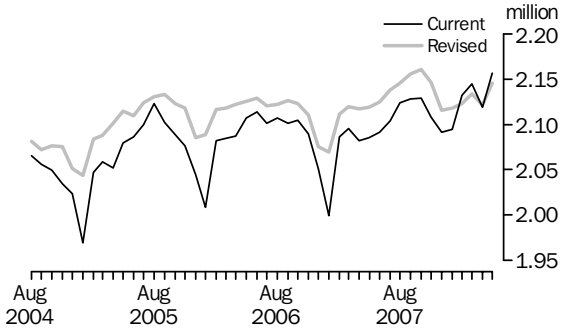
COUPLE FAMILIES, AUSTRALIA



COUPLE FAMILIES WITH CHILDREN UNDER 15 YEARS, AUSTRALIA

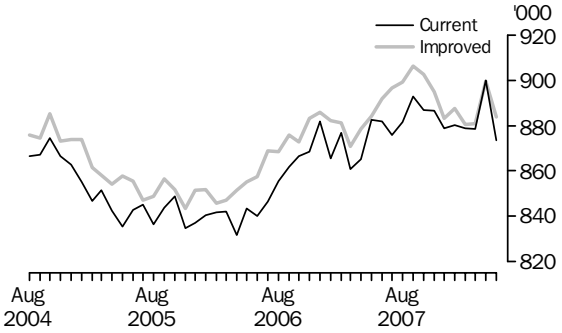


COUPLE FAMILIES WITH DEPENDANTS UNDER 25 YEARS, AUSTRALIA

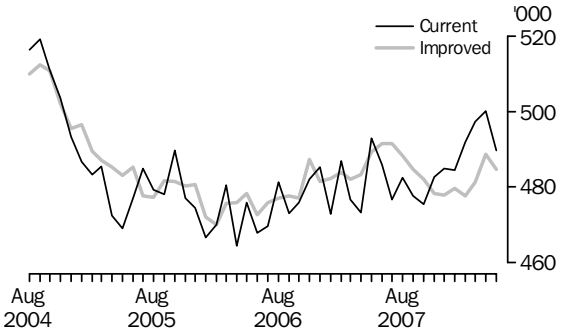


AUSTRALIA BY FAMILY TYPE *continued*

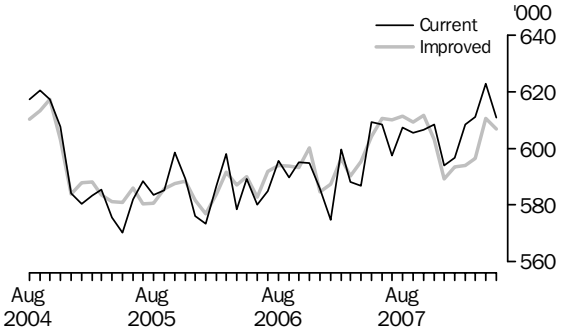
LONE PARENT FAMILIES, AUSTRALIA



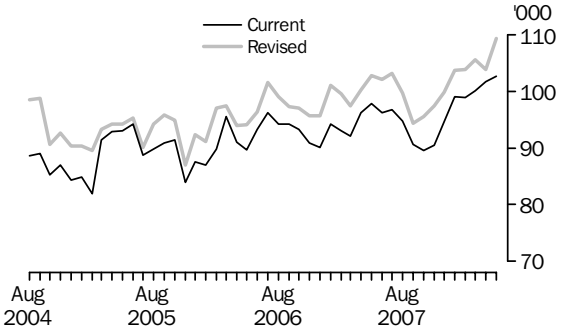
LONE PARENT FAMILIES WITH CHILDREN UNDER 15 YEARS, AUSTRALIA



LONE PARENT FAMILIES WITH DEPENDANTS UNDER 25 YEARS, AUSTRALIA

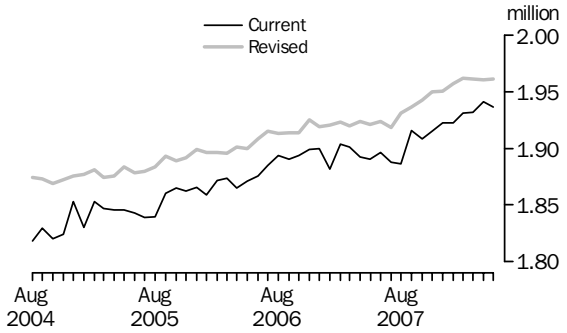


OTHER FAMILIES, AUSTRALIA

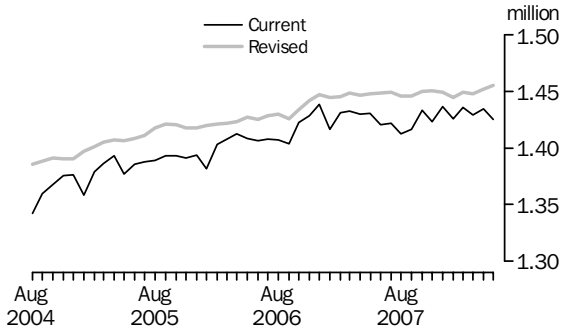


STATE/TERRITORY

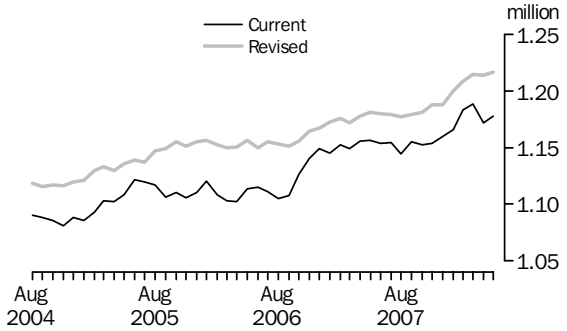
FAMILIES, NEW SOUTH WALES



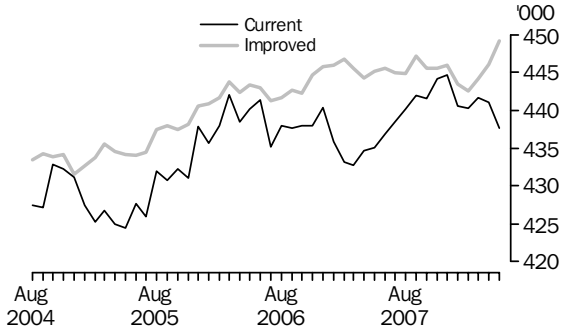
FAMILIES, VICTORIA



FAMILIES, QUEENSLAND

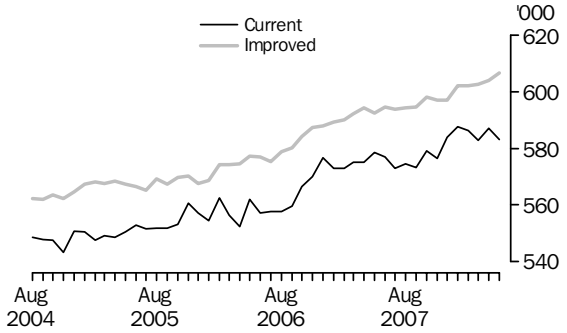


FAMILIES, SOUTH AUSTRALIA

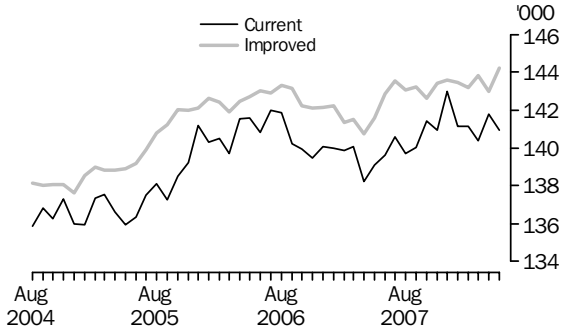


STATE/TERRITORY *continued*

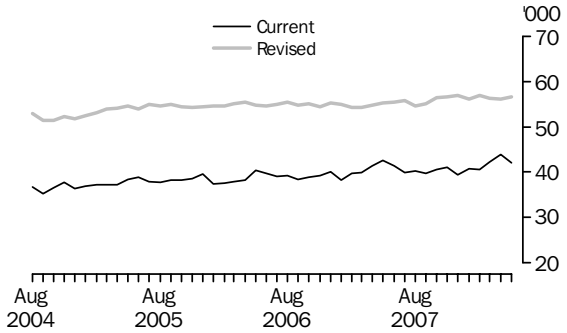
FAMILIES, WESTERN AUSTRALIA



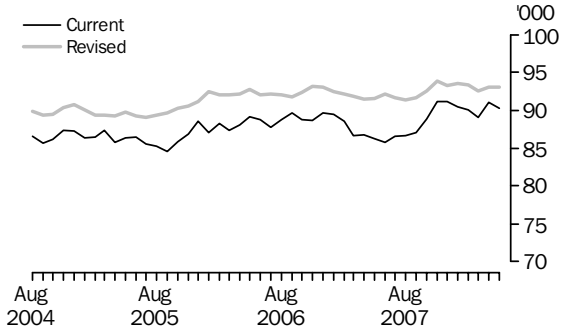
FAMILIES, TASMANIA



FAMILIES, NORTHERN TERRITORY



FAMILIES, AUSTRALIAN CAPITAL TERRITORY



GLOSSARY

Balance of State The aggregation of all Statistical Divisions within a State or Territory other than its Capital City Statistical Division. (Further details are included in Australian Standard Geographical Classification (ASGC) (cat. no. 1216.0)).

Benchmark An independent estimate to which survey responses are calibrated. For example, survey responses from the LFS are calibrated to independent estimates of the population aged 15 years and over, classified by geographic area, age and sex.

Computer Assisted Interviewing A method of data collection whereby responses are recorded directly into an electronic questionnaire on a notebook computer.

Datacube An interactive data file which can be cross-classified by a pre-determined set of variables. There are five LFS families datacubes, published under cat. no. 6224.0.55.001 on the ABS website in SuperTable format.

Dependants All family members under 15 years of age; family members aged 15–19 attending school or aged 15–24 attending a tertiary educational institution full time (except those classified as husbands, wives or lone parents).

Estimated Resident Population The official measure of the population of Australia based on the concept of residence. It refers to all people, regardless of nationality or citizenship, who usually live in Australia, with the exception of foreign diplomatic personnel and their families. It includes usual residents who are overseas for less than 12 months. It excludes overseas visitors who are in Australia for less than 12 months.

Family Two or more persons, 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 are usually resident in the same household. The basis of a family is formed by identifying the presence of a couple relationship, lone parent–child relationship or other blood relationship. Some households will, therefore, contain more than one family.

Family head Any person without a spouse present:

- with a son or daughter aged under 15 present; or
- without a son or daughter aged under 15 present, but with a son or daughter aged 15 or over present (provided that the son or daughter has no spouse, son or daughter of his/her own present).

Where a family has no person falling into either of these categories, the family head is generally defined to be the eldest person in the family.

No family head is determined for a couple family.

Family type In the LFS families datacubes, the categories for family type are:

1 Couple family

1.1 Couple family with dependants

1.1.1 Couple family with children under 15

1.1.2 Couple family without children under 15, but with dependent students

1.2 Couple family without dependants

1.2.1 Couple family without dependants, but with children 15 years or older

1.2.2 Couple family without children

2 Lone parent family

2.1 Lone parent family with dependants

2.1.1 Lone parent family with children under 15

2.1.2 Lone parent family without children under 15, but with dependent students

2.2 Lone parent family without dependants

3 Other families

Harmonic mean A method of calculating an average by dividing the number of observations by the sum of the reciprocals of each observed value. Under the current families estimation method, the harmonic mean is used to calculate the family weights from the person weights. For example, if a family consists of three people, and their person weights are 100, 200 and 300, the harmonic mean will be:

3 (the number of people in the family) divided by [1/100 + 1/200 + 1/300] = 164

Household A group of one or more persons in a private dwelling who consider themselves to be separate from other persons (if any) in the dwelling, and who make regular provision to take meals separately from other persons, i.e. at different times or in different rooms. Lodgers who receive accommodation but not meals are treated as separate households.

Boarders who receive both accommodation and meals are not treated as separate households. A household may consist of any number of families and non-family members.

Non Private dwelling An establishment which provides a communal type of accommodation, such as a hotel, motel, hospital or other institution.

Private dwelling A residential structure which is self-contained, owned or rented by the occupants, and intended solely for residential use. A private dwelling may be a flat, part of a house, or even a room, but can also be a house attached to, or rooms above shops or offices.

Relationship in household The relationship of all persons usually resident in a household to the household reference person. Where the relationship to the household reference person is other than a couple relationship or a parent-child relationship, a closer relationship to another household member is recorded, if one exists.

Sampling error Sampling error occurs because a sample, rather than the entire population, is surveyed. One measure of the likely difference resulting from not including all dwellings in the survey is given by the standard error. There are about two chances in three that a sample estimate will differ by less than one standard error from the figure that would have been obtained if all dwellings had been included in the survey, and about nineteen chances in twenty that the difference will be less than two standard errors.

Selection rules Rules applied in household surveys to ensure that each person is associated with only one dwelling, and hence has only one chance of selection.

Unit record data Data at the finest level of detail. For LFS, the finest level of detail is the person level. For confidentiality reasons, data are aggregated for output purposes.

Usual resident A person who usually lives in that particular dwelling and regards it as their own or main home.

Weights Factors applied to sample responses to expand them to produce population estimates.

FOR MORE INFORMATION . . .

INTERNET **www.abs.gov.au** the ABS website is the best place for data from our publications and information about the ABS.

INFORMATION AND REFERRAL SERVICE

Our consultants can help you access the full range of information published by the ABS that is available free of charge from our website. Information tailored to your needs can also be requested as a 'user pays' service. Specialists are on hand to help you with analytical or methodological advice.

PHONE 1300 135 070

EMAIL client.services@abs.gov.au

FAX 1300 135 211

POST Client Services, ABS, GPO Box 796, Sydney NSW 2001

FREE ACCESS TO STATISTICS

All statistics on the ABS website can be downloaded free of charge.

WEB ADDRESS www.abs.gov.au