6286.0.55.002

Australian Bureau of Statistics

Technical Paper

Australians' Employment and Unemployment Patterns: Expanded Confidentialised Unit Record File

Australia

1994 to 1997

100 years of STATISTICS

www.abs.gov.au

New Issue

Technical Paper

Australians' Employment and Unemployment Patterns: Expanded Confidentialised Unit Record File

Australia

1994 to 1997

Susan Linacre
Acting Australian Statistician

AUSTRALIAN BUREAU OF STATISTICS

EMBARGO: 11.30AM (CANBERRA TIME) MON 27 JUN 2005

ABS Catalogue No. 6286.0.55.002

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.

CONTENTS

ABBREVIATIONS

ABS	Australian Bureau of Statistics
CD	Collection District
CES	Commonwealth Employment Service
CURF	Confidentialised Unit Record File
DEET	Australian Government Department of Employment, Education and Training
DEETYA	Australian Government Department of Employment, Education, Training and Youth Affairs
DEWR	Australian Government Department of Employment and Workplace Relations
DSS	Australian Government Department of Social Security
FaCS	Australian Government Department of Family and Community Services
LFS	Labour Force Survey
LMP	Labour Market Program
PRG	Population Reference Group
RADL	Remote Access Data Laboratory
RSE	relative standard error
SE	standard error
SEUP	Survey of Employment and Unemployment Patterns

CHAPTER 1 – INTRODUCTION

OVERVIEW

This publication provides information about expanded Confidentialised Unit Record File (CURF) data from the 1994–1997 Survey of Employment and Unemployment Patterns (SEUP). Two microdata files are available from the survey:

- a basic CURF (Australians' Employment and Unemployment Patterns, 1994–97, Basic) available on CD-ROM; and
- an expanded CURF (Australians' Employment and Unemployment Patterns, 1994–97, Expanded) accessible only via the Remote Access Data Laboratory (RADL).

The RADL is an on-line database query system, under which microdata are held on a server at the ABS, to which users can submit programs to interrogate, analyse and model the data. Further information about this facility is available on the ABS web site <http://www.abs.gov.au> (Services We Provide - CURFs).

The sample for the SEUP included a list sample of known job seekers, provided by the then Department of Employment, Education and Training (DEET) (later known as the Department of Employment, Education, Training, and Youth Affairs (DEETYA), and now as the Department of Employment and Workplace Relations (DEWR)). Data collected directly from the respondent were supplemented with information about Commonwealth Employment Service (CES) registrations, case management and Labour Market Program participation from DEETYA and with data about income support from the (then) Department of Social Security (DSS) (now the Department of Family and Community Services (FaCS)). This only occurred with the respondent's consent. Throughout this paper the former department names are used in order to maintain continuity with data item names.

The DEETYA list sample of labour market program participants was not included on the CD-ROM Basic CURF, nor were the data items from the DEETYA and DSS administrative records, because of the danger of matching survey records to administrative data. Instead, a supplementary dataset of 500 records was provided to allow analyses against live data to be planned, and then subsequently run by the ABS. The supplementary dataset was manufactured from 'live' data, selected from the area sample component of the survey. Live DSS and DEETYA sourced data from statistically matched units were appended to the supplementary file units.

The Basic CURF contains fewer data items than the Expanded CURF and is for client use in their own computing environment (see the 'Comparison with the Basic CURF' section of Chapter 4 for information on the additional data items included in the Expanded CURF). There are no restrictions on the output that can be produced if the Basic CURF is accessed via CD-ROM.

This technical paper relates specifically to the Expanded CURF accessible through the RADL. The Expanded CURF includes supplementary administrative data about labour market program participation and CES registrations.

ABOUT THE SURVEY

The SEUP is a longitudinal survey. Data were collected from the same individuals (referred to as a panel) in three waves covering the period September 1994 to September 1997. The objectives of the survey were to provide information on the dynamics of the labour market and to help assess the impact of labour market assistance initiatives in alleviating joblessness in Australia.

ABOUT THE SURVEY *continued*

The panel was established in April to July 1995, for simplicity referred to as 'May 1995'. It included people who were aged 15–59 years and were living in private dwellings in both urban and rural areas. Initial data were also collected at this point. A further interview later in 1995 extended the time frame for which information was available to the full year ending September 1995, completing the collection of wave 1 data. Interviews to collect wave 2 and 3 data were undertaken during October 1996 (for the year ending September 1996) and October 1997 (for the year ending September 1997) respectively.

The three waves provided detailed labour market activity information, information on training and job offers, and a wide range of socio-demographic information for the period September 1994 to September 1997.

Primary data from the survey were released in the ABS publication *Australians' Employment and Unemployment Patterns 1994–1997* (cat. no. 6286.0). That publication is included as part of the CURF package.

The ABS has also published several Occasional Papers that use SEUP data. These papers present the results of analytical work done under a Research Fellowship scheme provided by the ABS. All of these papers are available on the ABS website and are listed below.

- Occasional Paper: *A Risk Index Approach to Unemployment – An Application Using the Survey of Employment and Unemployment Patterns* (cat. no. 6293.0.00.001)
- Occasional Paper: *Labour Market Programs, Unemployment and Employment Hazards* (cat. no. 6293.0.00.002)
- Occasional Paper: *Job Quality and Churning of the Pool of the Unemployed* (cat. no. 6293.0.00.003)
- Occasional Paper: *The Dynamics of Labour Market State and Benefit Receipt* (cat. no. 6293.0.00.004)
- Occasional Paper: *Labour Market Outcomes of Low Paid Adult Workers* (cat. no. 6293.0.00.005)
- Occasional Paper: *Labour Market Dynamics in Australia – An Application Using the Survey of Employment and Unemployment Patterns* (cat. no. 6293.0.00.006)
- Occasional Paper: *Dynamics of Earned Income in Australia – An Application Using the Survey of Employment and Unemployment Patterns* (cat. no. 6293.0.00.007)

CHAPTER 2 – SURVEY METHODOLOGY

SCOPE

The target population for the survey was those people considered to be most likely to be currently eligible for labour market assistance or to become eligible for assistance in the near future. This was determined after consultation with government officials, labour market analysts and other users.

The survey scope included all people resident in private dwellings who were aged 15–59 years, other than:

- overseas residents in Australia;
- certain diplomatic personnel of overseas governments, customarily excluded from the Census of Population and Housing and estimated resident population counts;
- members of non-Australian defence forces (and their dependants) stationed in Australia; and
- people living in remote areas that were difficult or costly to enumerate.

The exclusion of people living in remote areas has only a minor impact on aggregate estimates for individual states and territories, with the exception of the Northern Territory, where such people account for over 20% of the population.

Usual residents of selected private dwellings were included in the survey unless they were going to be absent from the dwelling until the end of the enumeration period. Other people present were considered to be visitors and were not included in the survey.

SAMPLE DESIGN

The sample was segregated into three subgroups: Jobseekers, people who were Labour Market Program (LMP) participants, and a Population Reference Group (PRG). All of the people in the sample were aged 15–59 years.

Jobseekers were the main component of the panel. The Jobseeker subgroup comprised those people considered most likely to be eligible to participate in a LMP or likely to become eligible for such assistance in the near future. Broadly speaking, Jobseekers were people who, at May 1995, were in one of the following categories:

- unemployed;
- marginally attached to the labour force (e.g. discouraged jobseekers); or
- underemployed.

The LMP subgroup was a sample of people who had been in continuous receipt of unemployment allowance for 18 months and who had started a subsidised employment placement and/or a labour market training program between July 1994 and February 1995. This component of the panel complemented the Jobseekers subgroup, and was included to ensure that the survey had enough LMP participants to support analysis of their characteristics in the first year.

The PRG was a random sample of the population aged 15–59 years. It was included in the panel so that the labour market experience of Jobseekers could be assessed in the context of the experience of the general population.

SAMPLE DESIGN *continued*

With respondents' consent, data collected directly from them during the interview were supplemented with data about CES registrations, case management and LMP participation (from DEETYA) and data about income support (from DSS). This minimised the interview time for respondents and ensured that accurate information was available about their involvement with labour market assistance programs and about their receipt of income support.

A probability sample design was used. The Jobseeker and Population Reference Group samples were drawn from the Population Survey Master Sample and excluded non-private dwellings (e.g. hotels) and the remote and sparsely settled stratum. The Labour Market Program participants sample was drawn from a list of such participants provided by DEETYA.

Jobseeker subgroup

For the Jobseeker subgroup, three major adjustments were made to the usual population survey sample selection methodology:

- Census Collection Districts (CDs) were selected in the survey sample with unequal probabilities according to the number of unemployed people resident in the CD at the time of the 1991 Census enumeration. CDs with a high proportion of resident unemployed people in August 1991 were given a higher probability of selection in the survey sample than CDs with a low proportion of unemployed people. The weighting scheme for the Jobseeker subgroup accounted for this unequal probability sampling scheme.
- More dwellings than usual were selected in the survey sample to be screened for Jobseekers – approximately twice as many dwellings as were in the Labour Force Survey sample in 1994–1997.
- The sample was designed to be more geographically clustered than the Labour Force Survey. Such clustering produced a substantial saving in survey field costs for only a small loss of sample efficiency.

For the Jobseeker subgroup, all people aged 15–59 years in selected dwellings were screened to determine whether or not they were a 'Jobseeker'.

The Jobseeker subgroup was identified in May 1995, whereas the reference period for wave 1 of the Expanded CURF starts at September 1994. Thus at the beginning of wave 1 there were some person records classified as belonging to the Jobseeker subgroup that did not meet the criteria for classifying Jobseekers. For example a respondent could have been working at the start of wave 1, but be classed as a Jobseeker because at some time in May 1995 they were unemployed.

Labour Market Program participants

The sample methodology used for this subgroup was a list-based probability sample. DEETYA provided the ABS with a list of eligible people. In order to produce a more efficient sample, the list was sorted into similar groups (strata) and then a random sample was selected from each group. The stratification variables used were: state of usual residence, length of unemployment, and job placement/training status. In total 2,300 people were selected. This number allowed for various types of sample loss, such as people refusing to allow DEETYA to pass their name to the ABS, incorrect addresses and overlap with the Labour Force Survey sample.

Labour Market Program participants continued

Procedures were put in place to ensure that the probability sample chosen for this subgroup was geographically close to the Jobseeker subgroup and the Population Reference Group, but did not overlap these subgroups or any other ABS household survey.

Population Reference Group

The PRG subgroup was a random sample of the population aged 15–59 years. It was included so that outcomes of labour market programs could be assessed in the context of general labour market conditions and movements. It was also a source of longitudinal information for the general population on other topics covered by the survey, such as housing and income.

A random subsample of the SEUP dwelling sample was chosen to provide the PRG dwelling sample, with one person then selected from each PRG dwelling to create the PRG subgroup.

Using this sampling methodology, it was possible for a person to be selected both in the Jobseeker subgroup and the Population Reference Group.

COVERAGE RULES

Coverage rules were applied to the Jobseeker and PRG subgroups, to ensure that each person in scope was associated with only one dwelling and hence had only one chance of selection. Coverage rules were not needed for the LMP subgroup as specific people had already been identified as members of this subgroup.

PANEL SIZE AND MAINTENANCE

The ability to maintain contact with a relatively high proportion of the panel was critical to the usefulness of the survey data. A number of strategies were used to achieve this, including:

- providing 'change of details' cards for respondents to advise new contact details, etc.;
- providing a toll-free telephone number for respondents to call;
- asking the respondent for contact details of up to three people who were likely to know the respondent's whereabouts; and
- regular mail contact throughout the survey.

However, it was inevitable that some non-response would occur when people were unwilling or unable to cooperate, or when they could not be contacted.

The attrition rate is the percentage of previous wave respondents who did not respond in the current wave. Attrition between waves caused a permanent drop in the sample size as the survey did not replace non-respondents. Although the weighting procedure for each wave partly corrected for attrition in the sample, there are some small differences in estimates for each wave. Higher than average sample loss occurred for males, young people, and people who were renting accommodation. Table 2.1 shows the size and composition of the panel for each wave and the overall attrition rate.

2.1 COMPOSITION AND SIZE OF THE PANEL, People

Subgroup	*Wave 1* no.	*Wave 2* no.	*Wave 3* no.	*Attrition rate (Wave 1 – Wave 3)* %
Jobseekers(a)	5 488	4 779	4 261	22.4
Labour Market Program participants	1 019	888	775	23.9
Population Reference Group	2 311	2 120	1 983	14.2
Total panel	**8 591**	**7 585**	**7 019**	**18.3**

(a) Includes some Jobseekers who were also part of the Population Reference Group.

ESTIMATION PROCEDURE

Estimates from SEUP were calculated using a ratio estimation procedure, which ensured that the survey estimates relating to the Jobseeker population and to the Population Reference Group conform to independently estimated distributions (benchmarks) of these populations by age and sex, rather than to the age and sex distribution within the sample itself.

RELIABILITY OF ESTIMATES

Two types of error are possible in an estimate based on a sample survey; sampling error and non-sampling error. The sampling error is a measure of the variability that occurs by chance because a sample rather than an entire population is surveyed.

Since estimates from SEUP data are based on information obtained from a sample of people, they are subject to sampling variability; that is, they may differ from the estimates that would have been produced if all in-scope people had been included in the survey. One measure of the likely difference is given by the standard error, which indicates the extent to which an estimate might have varied by chance because only a sample of people was included. There are about two chances in three that a sample estimate will differ by less than one standard error from the estimate that would have been obtained if all people had been included, and about nineteen chances in twenty that the difference will be less than two standard errors.

Another measure of sampling variability is the relative standard error, which is obtained by expressing the standard error as a percentage of the estimate to which it refers. The relative standard error is a useful measure in that it provides an immediate indication of the percentage errors likely to have occurred due to sampling, and thus avoids the need to also refer to the size of the estimate.

The imprecision due to sampling variability should not be confused with inaccuracies that may occur because of imperfections in reporting by respondents, errors made in collection such as in recording and coding data, and errors made in processing the data. Inaccuracies of this kind are referred to as the non-sampling error and can occur in any enumeration, whether it be a full count or a sample. It is not possible to quantify non-sampling error, but every effort is made to reduce it to a minimum.

Appendix 1 provides information on standard errors for SEUP.

CHAPTER 3 – STRUCTURE OF THE EXPANDED CURF

RECORD TYPES

There are two types of record: person level records and episode level records. These two record types can be distinguished from each other using the level identifier as follows:

- person level records have a person level identifier (IDP) = 4;
- episode level records have an episode level identifier (IDQ) = 5.

A respondent's person record is linked to their associated episode record via the person number, i.e. person records and associated episode records share the same person number. For both person and episode level records, the person number has the field name ABSPID.

At the person level, records include socio-demographic data about the respondent, such as age, sex and marital status, labour force status, income and educational attainment. There are also data items for data collected at the episode level.

An 'episode' is defined as a period of time during which a particular activity is undertaken (such as working, looking for work or absence from the labour market). An episode can occur wholly within a reference period or can span more than one reference period.

The episode level comprises the following episodes and occurrences:

- episodes of Labour Market Activity, i.e. episodes of work, looking for work and absence from the labour market;
- episodes of Labour Market Support from DEETYA and DSS; and
- training and job offer occurrences.

A complete description of person and episode data items is included in Appendix 4 and Appendix 5. These appendices are Excel spreadsheets and are available from the ABS web site. Follow the link to [AusStats - Publications and Data], then [Data cubes]. They are listed as Appendix 4 and Appendix 5 under catalogue number 6286.0.55.002.

Person level records

Person records are uniquely identified via the person number (ABSPID).

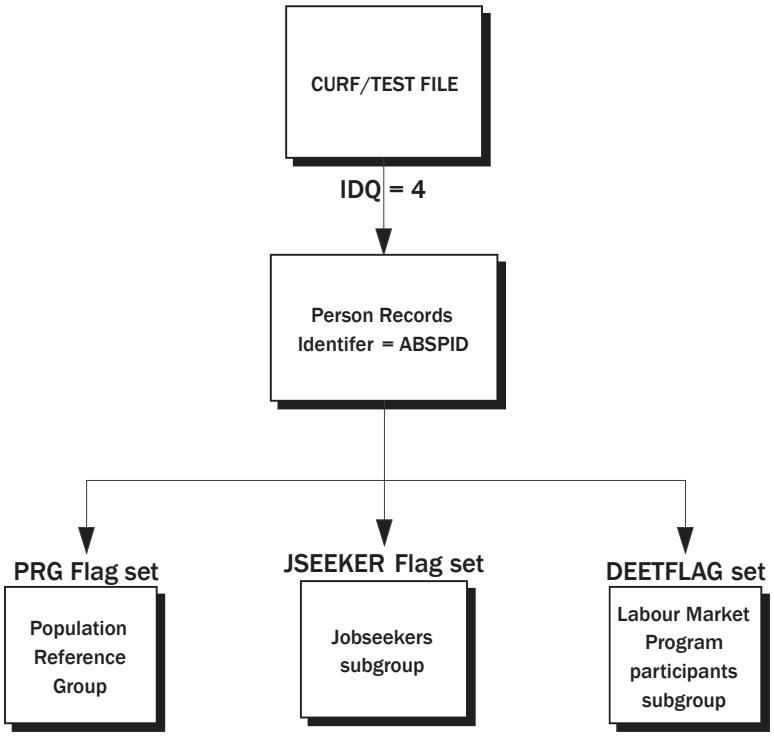
The following flags on a person record are used to classify it to one or more of the following subgroups:

- JSEEKER - identifies Jobseekers;
- PRG - identifies people from the Population Reference Group; and
- DEETFLAG - identifies people who are in the LMP subgroup.

Records on the Expanded CURF that belong to both the Jobseeker subgroup and the PRG have both the JSEEKER and PRG flag set.

Person level records continued

3.1 Subgroups for Person Level Records



Episode level records

Associated with each person are one or more episode and occurrence records. These records provide a detailed picture of each respondent's labour market activities. They also include details of job offers, training courses and income support from DSS, and details of case management, labour market program participation and CES registration from DEETYA.

The episode status field (EPSTAT), which describes the particular type of episode or occurrence, can assume the following values.

Episodes of Labour Market Activity

- EPSTAT = 1 identifies episodes of work
- EPSTAT = 2 identifies episodes of looking for work
- EPSTAT = 3 identifies episodes of absence from the labour market

Episodes of Labour Market Support

- EPSTAT = 4 Identifies episodes of DEETYA support. These episodes are further classified via the data item RECTYP as follows:
  - RECTYP = 1 identifies episodes of CES registration
  - RECTYP = 2 identifies episodes of case management
  - RECTYP = 3 identifies episodes of Labour Market Program Participation
- EPSTAT = 5 identifies episodes of Income Support from DSS

Occurrences

- EPSTAT = 6 identifies In-house Training Occurrences
- EPSTAT = 7 identifies External Training Occurrences
- EPSTAT = 8 identifies Job Offer Occurrences

Episode level records continued

It is possible for there to be multiple instances of each episode type for a respondent. For example during wave 2 a respondent could have had two job offers and three working episodes. For this reason each episode and occurrence is assigned an episode number (ABSEID).

For episodes of labour market activity, episode numbers are allocated based on each episode's beginning date (BEGDATE) and are dependent on the episode numbers of other episodes of labour market activity. For example, if the earliest episode of labour market activity for a respondent is a working episode then this is given an episode number of 1, if the next episode is an episode of looking for work then this is given an episode number of 2 etc.

For episodes of DEETYA and DSS support, episode numbers are also allocated based on BEGDATE, however each episode type is numbered independently of other episode types. For example the earliest episode of DEETYA data will be given an episode number of 1, as will the earliest episode of income support.

Episode numbers for occurrences are allocated the same way as they are for episodes of DEETYA and DSS support. However the data item BEGDATE refers to the date of the occurrence, for example the date a job offer was received or the date a training course commenced, rather than the start date of an episode.

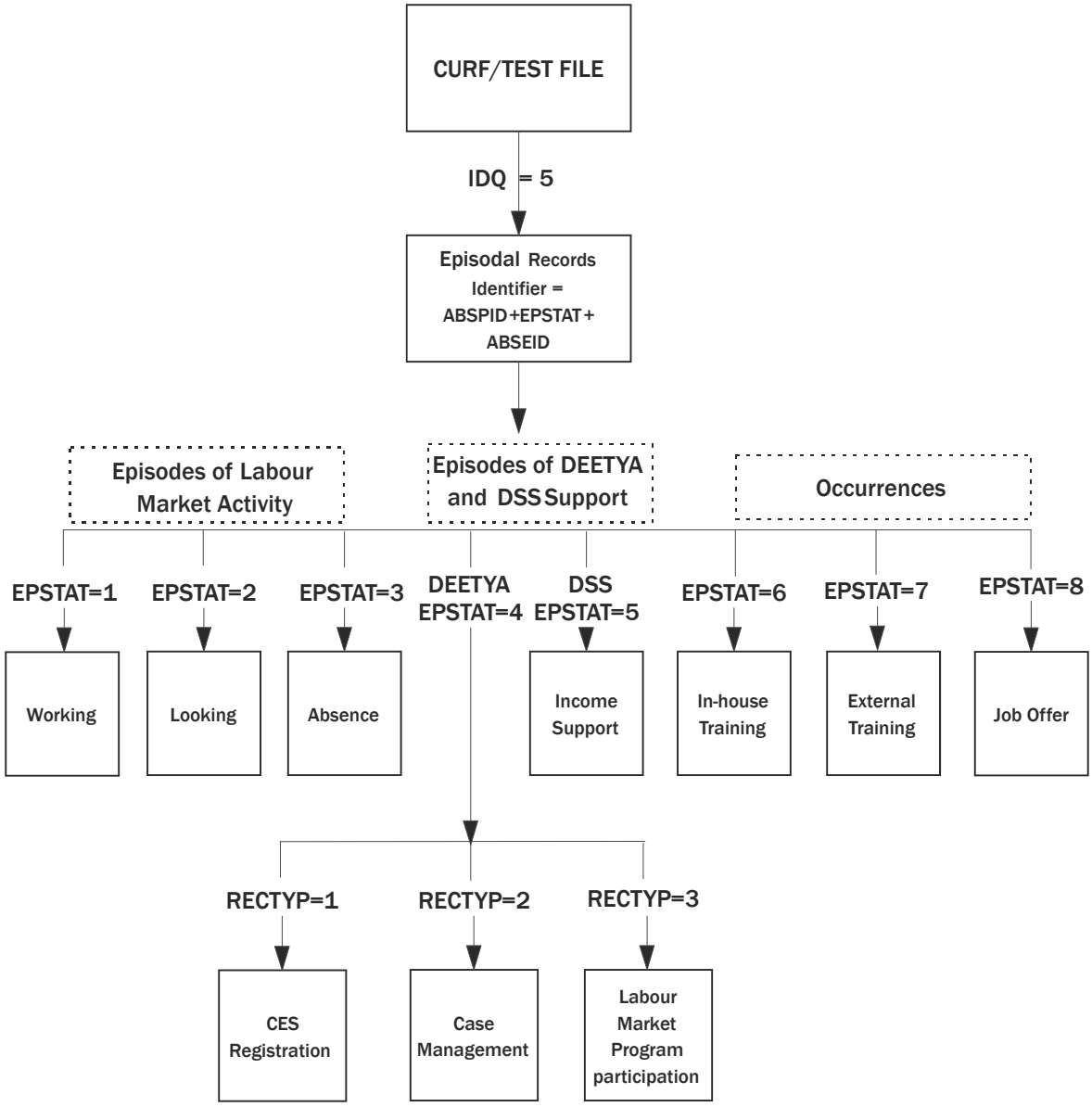
Due to the different numbering schemes for episode number, a person number (ABSPID), episode number (ABSEID) and episode status (EPSTAT) are required to uniquely identify an episode or occurrence on the Expanded CURF (see figure 3.2).

Although the record structure at the episode level is exactly the same for each type of episode, each episode or occurrence only contains data relevant to that particular type of episode or occurrence. For example an episode of working only contains information for data items such as hours worked, occupation and industry. If the respondent also had a period of looking for work during the reference period then a separate episode of looking, describing difficulties in finding work, steps taken to find work etc. would exist on the Expanded CURF.

Episodes can span any or all of the reference period(s). The data items BEGDATE and ENDDATE indicate the dates covered by the episode. Note, episodes commencing before the start of wave 1, (i.e. with BEGDATE prior to 5 September 1994) are in scope of the survey provided they are still continuing at the start of wave 1 (i.e. their ENDDATE is on or after 5 September 1994). If an episode is still continuing at the end of wave 3 it will have the flag CONTIN3 set and an ENDDATE of 31 August 1997.

Every day of the reference period should be covered by at least an episode of working, an episode of looking for work, or an episode of absence from the labour market. It is also possible for a respondent to have concurrent episodes of working, and episodes of working concurrent with episodes of looking. It is not possible to have concurrent episodes of looking nor is it possible to have an episode of absence concurrent with an episode of working, an episode of looking or another episode of absence.

3.2 Episodes and Occurrences on the CURF and TEST files



CHAPTER 4 – USING THE CURF DATA

ABOUT THE DATA

The SEUP data are released under the *Census and Statistics Act 1905*. The Act provides for the release of data in the form of unit records where the information is not likely to enable the identification of a particular person or organisation. Accordingly, there are no names and addresses of survey respondents on the CURF.

Intending purchasers should ensure that the data they need, at the level of detail they require, are available on the Expanded CURF. Appendix 4 and Appendix 5 contain a full list of available person and episode data items. These appendices are Excel spreadsheets and are available from the ABS web site. Follow the link to [AusStats - Publications and Data], then [Data cubes]. They are listed as Appendix 4 and Appendix 5 under catalogue number 6286.0.55.002.

There are 7,572 confidentialised respondent records available for analysis from SEUP. Subject to the limitations of the sample size and the data classifications used, it is possible to manipulate the data, produce tabulations and undertake statistical analyses to individual specifications.

Steps to confidentialise datasets made available on the Expanded CURF ensure the integrity of the dataset and optimise its content, while maintaining confidentiality of respondents. The steps taken to preserve confidentiality may include:

- reducing the level of detail for some items (e.g. geographic and demographic);
- ranging or collapsing the values of some variables; and
- modifying some records identified as high-risk.

As a result of these changes, it may not always be possible to reconcile the data produced from the Expanded CURF with published data.

CURF CONTENTS

The Expanded CURF can only be accessed via the RADL using SAS or SPSS at this time. STATA is expected to be available by late 2005.

The Expanded CURF contains information files, test files and main files as listed below.

Information files

DATA ITEM AND FREQUENCY LISTING.TXT – documentation on the data files including data item labels, field start positions and lengths, code values, category labels and frequency counts.

6286055002_1994 to 1997.PDF – an Acrobat file that contains this paper.

RESPONSIBLE ACCESS TO CURFS.PDF – an Acrobat file that explains CURF users' roles and obligations when using confidentialised data.

62860_1994 to 1997.PDF – an Acrobat file that contains the publication *Australians' Employment and Unemployment Patterns 1994-1997* (cat. no. 6286.0).

Test files

The test files mirror the structure of the actual data files but contain a smaller set of random data and identifiers. The test files are located on the RADL and can be used to trouble shoot code before submitting RADL jobs.

FORMATS.SC2 – a SAS library containing formats for the test files.

Test files continued

SEU97psn.SD2 – the test file for person level data for the Expanded CURF in SAS for Windows format.

SEU97eps.SD2 – the test file for episode level data for the Expanded CURF in SAS for Windows format.

SEU97psn.SAV – the test file for person level data for the Expanded CURF in SPSS format.

SEU97eps.SAV – the test file for episode level data for the Expanded CURF in SPSS format.

SEU97psn.DTA[1] – the test file for person level data for the Expanded CURF in STATA format.

SEU97eps.DTA[1] – the test file for episode level data for the Expanded CURF in STATA format.

Main files

FORMATS.SC2 – a SAS library containing formats.

PERSON.SD2 – the person level data for the Expanded CURF in SAS for Windows format.

EPISODE.SD2 – the episode level data for the Expanded CURF in SAS for Windows format.

PERSON.SAV – the person level data for the Expanded CURF in SPSS format.

EPISODE.SAV – the episode level data for the Expanded CURF in SPSS format.

PERSON.DTA[1] – the person level data for the Expanded CURF in STATA format.

EPISODE.DTA[1] – the episode level data for the Expanded CURF in STATA format.

COMPARISON WITH THE BASIC CURF

The Expanded CURF contains more completed variables than the Basic CURF. In the Basic CURF, data items from DEETYA and DSS administrative data were set to zero, however the Expanded CURF contains complete data for these items. A complete list of the CURF data items is included in Appendix 4 (Persons) and Appendix 5 (Episodes). The following tables show the additional data items available in the Expanded CURF.

4.1 DSS ADMINISTRATIVE DATA ITEMS

	IDENTIFIER		
Data item label	*Wave 1*	*Wave 2*	*Wave 3*
Total amount of income support in dollars (from DSS data)	TOTIS1	TOTIS2	TOTIS3
Total duration of income support in each wave for unemployment benefit	GISTDW11	GISTDW12	GISTDW13
Whether received income support during reference period	WHRIS1	WHRIS2	WHRIS3
Whether received income support (family payments) during each wave	RECINC31	RECINC32	RECINC33
Whether received income support (other allowances) during each wave	RECINC41	RECINC42	RECINC43
Whether received income support (pensions) during each wave	RECINC21	RECINC22	RECINC23
Whether received income support (unemployment benefits) during each wave	RECINC11	RECINC12	RECINC13

1 The 1994 to 1997 SEUP expanded CURF can only be accessed in SAS and SPSS at this time. STATA is expected to be available by late 2005.

4.2 DEETYA ADMINISTRATIVE DATA ITEMS

	IDENTIFIER		
Data item label	*Wave 1*	*Wave 2*	*Wave 3*
English skills (Reading)	READPROF	READPROF	READPROF
English skills (Speaking)	SPKPROF	SPKPROF	SPKPROF
English skills (Written)	WRITPROF	WRITPROF	WRITPROF
Total duration of case management in reference period	DURCAS1	DURCAS2	DURCAS3
Total duration of CES registration in reference period	DURCES1	DURCES2	DURCES3
Total duration of LMPS in reference period	DURLMP1	DURLMP2	DURLMP3
Whether participated in a brokered or other employment program during the reference period	WHBRO1	WHBRO2	WHBRO3
Whether participated in a job club during the reference period	WHJC1	WHJC2	WHJC3
Whether participated in a training LMP during the reference period	WHETP1	WHETP2	WHETP3
Whether participated in a wage subsidy program during the reference period	WHWS1	WHWS2	WHWS3
Whether participated in an apprenticeship/traineeship LMP during the reference period	WHAPP1	WHAPP2	WHAPP3
Whether participated in an employment support program during the reference period	WHES1	WHES2	WHES3
Whether participated in labour market program during reference period	WHLMP1	WHLMP2	WHLMP3
Whether received case management during reference period	WHECAS1	WHECAS2	WHECAS3
Whether registered with CES during reference period	WHECES1	WHECES2	WHECES3

DATA ITEMS

Data items included in the Expanded CURF are listed in Appendix 4 (Persons) and Appendix 5 (Episodes). These appendices are Excel spreadsheets and are available from the ABS web site. Follow the link to [AusStats - Publications and Data], then [Data cubes]. They are listed as Appendix 4 and Appendix 5 under catalogue number 6286.0.55.002.

There are two types of data items on the Expanded CURF: categorical variables and continuous variables.

A categorical variable is one where the responses belong to a number of discrete categories. For example 'Status in Employment' (EMPSTAW) has the categories Employee, Employer, Own account worker and Contributing family worker.

A continuous variable is one where the response does not form part of a set of discrete categories. Instead the responses may take any form within the limits of the range of characters or numbers available for that particular variable. For example, the continuous variable 'Duration of employment of last full-time job' (EMPDURFT) can have responses anywhere in the range of 0 to 9999 months.

To help analyse the Expanded CURF, data items that are continuous are presented in the documentation file DATA ITEM AND FREQUENCY LISTING.TXT in a grouped format. For example, although the variable EMPDURFT may have values such as 1 month, 15 months, or 75 months, the relevant groups for those responses would be Under 2 months, 13–18 months, Over 5 years etc.

Continuous data items also have special codes to represent responses that could not normally be represented by a number. For example for the data item YLFTES (Years since left full time education), a code of 96 indicates that the respondent is still studying, and a code of 95 indicates that they finished studying less than 1 year previously. A full listing of these special codes is in Appendix 3.

Fixed data

Fixed data at the Person level are those that were collected in wave 1 and are not subject to change over time. They include sex, birthplace and industry of last full time job (prior to 6 September 1994). Fixed data at the Episode level could be collected in wave 1 or wave 2 and not subject to change over time. Examples of fixed data items at the episode level include an episode's beginning date, industry and labour market activity of the episode.

Dynamic data

Dynamic data may be subject to change over time, for example marital status and income at the person level, or hours worked and full time/part time status of a job at the episode level.

If a respondent participated only in wave 1 of the survey then person level and episode level dynamic data items are set to not applicable for wave 2 and wave 3. If a respondent participated in all waves of the survey, but their working episode only covered waves 2 and 3, then dynamic data items for the working episode will be set only for waves 2 and 3.

If a dynamic data item has been collected for more than one wave, analysts will need to determine an appropriate approach to the analysis.

Note, because occurrences describe a training course or job offer at a point in time, none of the data items relating to occurrences are dynamic.

IMPUTATION

Wave non-response occurred when respondents did not participate in survey waves. The weighting algorithm used in deriving estimates allowed for non-response.

Item non-response occurred where a respondent either did not know or refused to answer a particular survey question. Imputation was used to replace these missing values for selected data items. Overall, just over half of the person records and about two percent of the episode records had some degree of imputation. However, only a relatively small number of items were selected for imputation.

The two methods of imputation used in the SEUP were cross wave imputation and imputation within a wave.

In cross wave imputation, a missing data item for a wave was replaced with earlier wave data for the same record. When imputing data, account was taken of average movements in the data item between waves. For example if the data item increased by an average of 5% between wave 1 and wave 2, then the data item imputed for wave 2 would be 5% greater than the corresponding data item for wave 1.

Imputation within a wave involved imputing the missing data from a record using a 'donor' record identified as having similar characteristics.

Each record on the Expanded CURF has a series of imputation flags - one for each data item potentially subject to imputation. If the data item was imputed then the corresponding imputation flag is set. These flags allow analysts to include/exclude imputed records from their analyses, or to do their own imputation.

WEIGHTING AND ESTIMATION

A number of weights are included in the Expanded CURF for the different population groups. These weights are longitudinal and account for attrition rate over subsequent waves. For example, when applied to wave 2 person level data items on the Expanded CURF, the estimates of total populations for a given subgroup will be the same as those at wave 1, i.e. 875,100 Jobseekers and 11,050,500 General Population.

Weights are stored in the SEUP data files without a decimal point. This permits the inclusion of a larger set of digits within the fixed range of numeric characters available for each of the weights variables, and results in a greater degree of accuracy in the production of weighted data. *Consequently all weights should be divided by a factor of 10,000 before being applied to the data.*

Standard weights

Standard weights should be used if the analysis does not use data items which are sourced from DEETYA and DSS (see below). The following table shows the standard weights that are available, their data item name at the Person and Episode level, the subgroup to which they apply and the number of respondents in each subgroup. For example the PRG weight for wave 1 (PRGWGT1) applies to the 2,311 respondents in the PRG who participated in wave 1, whereas because of attrition the PRG weight for wave 2 applies to only 2,120 people.

4.3 STANDARD WEIGHTS FOR EACH SUBGROUP

	DATA ITEM NAME		
Weight	*Person level*	*Episode level*	*Subgroup size*
Population Reference Group			
PRG Weight - wave 1	PRGWGT1	PRGWGTQ1	2 311
PRG Weight - wave 2	PRGWGT2	PRGWGTQ2	2 120
PRG Weight - wave 3	PRGWGT3	PRGWGTQ3	1 983
Jobseekers			
Jobseeker Weight - wave 1	JSWGT1	JSWGTQ1	5 488
Jobseeker Weight - wave 2	JSWGT2	JSWGTQ2	4 779
Jobseeker Weight - wave 3	JSWGT3	JSWGTQ3	4 261
Labour Market Program Participants			
LMP Weight - wave 1	LMPWGT1	LMPWGQ1	1 019
LMP Weight - wave 2	LMPWGT2	LMPWGQ2	888
LMP Weight - wave 3	LMPWGT3	LMPWGQ3	775

The choice of a particular standard weight depends on the subgroup and the wave being analysed. For example if analysis requires wave 2 data for Jobseekers, and there are no administrative data items included in the analysis, the weight JSWGT2 should be used.

DEETYA and DSS weights

Data from the administrative systems of DEETYA and DSS were only obtained for respondents who gave their consent. Table 4.4 below shows the DEETYA weights, DSS weights and combined DEETYA/DSS weights for waves 2 and 3.

Due to difficulties with the administrative data received for wave 1, the administrative data was not applied to the overall wave 1 SEUP data. Administrative data used in wave 2 were a combination of wave 2 data and wave 1 data, and use wave 2 weights. Wave 3 incorporated administrative data from all three waves and uses wave 3 weights.

DEETYA and DSS weights continued

DEETYA weights should be used for data items sourced from that Department, i.e. data on Labour Market Program participation. DSS weights should be used when using DSS data items, i.e. data on Income Support. Combined weights should be used if both DEETYA and DSS data items are to be used in the same analysis. For example, if the analysis of wave 2 Jobseekers episodes includes data items from DSS then the weight DSJSWTQ2 should be used. If administrative data from both DEETYA and DSS was part of the analysis then weight DSDTJSQ2 would be appropriate.

4.4 DEETYA AND DSS WEIGHTS FOR EACH SUBGROUP

	DATA ITEM NAME		
Weight	*Person level*	*Episode level*	*Subgroup size*
Jobseekers consenting to provide DSS data			
DSS Jobseeker Weight - wave 2	DSJSWT2	DSJSWTQ2	3 958
DSS Jobseeker Weight - wave 3	DSJSWT3	DSJSWTQ3	3 613
Persons from PRG consenting to provide DSS data			
DSS PRG Weight - wave 2	DSPRGWT2	DSPRGWQ2	1 020
DSS PRG Weight - wave 3	DSPRGWT3	DSPRGWQ3	1 005
Labour Market Program Participants consenting to provide DSS data			
DSS LMP Weight - wave 2	DSLMWT2	DSLMWTQ2	797
DSS LMP Weight - wave 3	DSLMWT3	DSLMWTQ3	701
Jobseekers consenting to provide DEETYA data			
DEETYA Jobseeker Weight - wave 2	DTJSWT2	DTJSWTQ2	3 951
DEETYA Jobseeker Weight - wave 3	DTJSWT3	DTYJSWTQ3	3 580
Persons from PRG consenting to provide DEETYA data			
DEETYA PRG Weight - wave 2	DTPRGWT2	DTPRGWQ2	823
DEETYA PRG Weight - wave 3	DTPRGWT3	DTPRGWQ3	797
Labour Market Program Participants consenting to provide DEETYA data			
DEETYA LMP Weight - wave 2	DTLMWT2	DTLMWTQ2	822
DEETYA LMP Weight - wave 3	DTLMWT3	DTLMWTQ3	724
Jobseekers consenting to provide DSS and DEETYA data			
DSS/DEETYA combined Jobseeker Weight - wave 2	DSDTJSW2	DSDTJSQ2	3 725
DSS/DEETYA combined Jobseeker Weight - wave 3	DSDTJSW3	DSDTJSQ3	3 385
Persons from PRG consenting to provide DSS and DEETYA data			
DSS/DEETYA combined PRG Weight - wave2	DSDTPRW2	DSDTPRQ2	771
DSS/DEETYA combined PRG Weight - wave 3	DSDTPRW3	DSDTPRQ3	745
Persons from PRG consenting to provide DSS and DEETYA data			
DSS/DEETYA combined LMP Weight - wave 2	DSDTLMW2	DSDTLMQ2	790
DSS/DEETYA combined LMP Weight - wave 3	DSDTLMW3	DSDTLMQ3	697

Not all respondents provided consent for the ABS to obtain relevant DSS and/or DEETYA data. For example, although there were 2,120 people in the PRG subgroup for wave 2, only 823 people from this subgroup agreed to the use of DEETYA data. Records for non-consenting units will have a DSS and/or DEETYA weight of 0 and a combined DSS/DEETYA weight of zero. For example, a Jobseeker who agreed to allow DSS to provide data but did not provide consent for DEETYA data will have a DSS Jobseeker weight but will not have a DEETYA Jobseeker weight nor a DSS/DEETYA combined Jobseeker weight.

DEETYA and DSS weights continued

For the above reasons, the estimates obtained using DEETYA and DSS weights will differ somewhat from those obtained using standard weights. For example using the standard Jobseeker weight JSWGT2, 98,900 Jobseekers had an occupation of Labourer or related worker, whereas using the non standard DEETYA weight DTJSWT2 103,100 Jobseekers had this occupation.

Choosing an appropriate weight

Users should take care in selecting an appropriate weight to use in analysis. If an inappropriate weight is used, this can have a significant effect on the results of the analysis. For example:

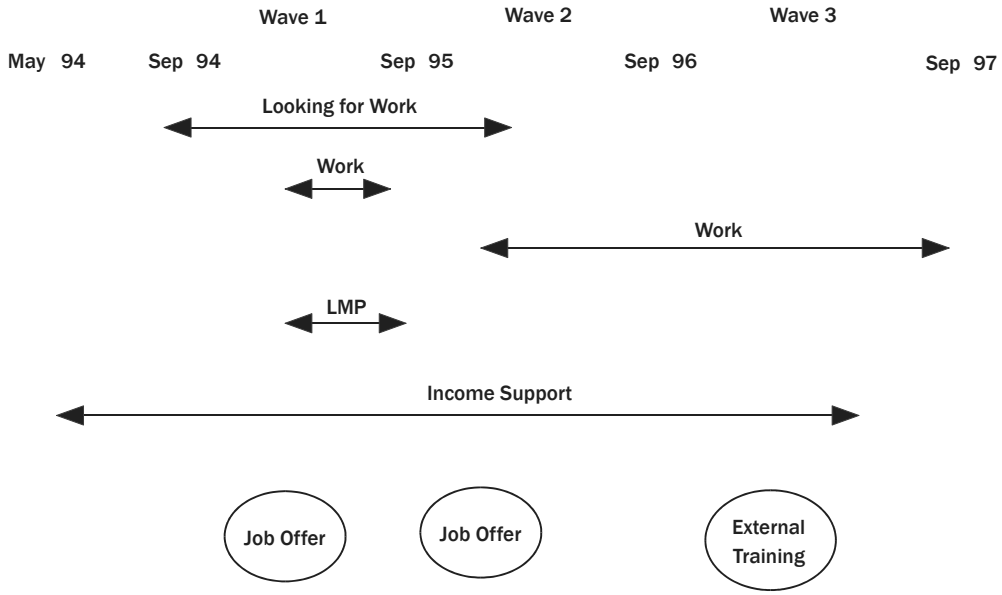
- if administrative data is included in the analysis but the weight JSWGT1 is used then the results will include a high rate of non-availability for the administrative data items. This will be from people who did not consent to the provision of DSS and DEETYA data; and
- if administrative data is not included in the analysis but a non-standard weight such as DSDTJSW2 is used, then the estimates produced will have a higher standard error than if the appropriate weight, JSWGT2 is used. The higher standard error occurs because the weight DSDTJSW2 is based on a smaller sample of records.

Details of standard errors for weighted estimates are included in Appendix 1.

ANALYSING EPISODES AND OCCURRENCES

The following diagram shows a series of episodes and occurrences for a hypothetical respondent.

4.5 Activities for a hypothetical respondent



The data item CONTIN2 is used to identify episodes that are still continuing at the end of wave 2. Episodes that were completed before the start of wave 3 have CONTIN2 set to No. Table 4.6 shows the treatment of these episodes and occurrences.

4.6 EPISODES AND OCCURRENCES, Hypothetical respondent

Episode status (EPSTAT)	*EPINUM*	*BEGDATE*	*ENDDATE*	*CONTIN2*
2 (Looking for work)	1	1 Apr 94	1 Jan 96	No
1 (Working)	2	1 Mar 95	1 Jun 95	No
1 (Working)	3	1 Oct 95	31 Aug 97	Yes
4 (LMP)	1	1 Apr 95	1 Jul 95	. .
5 (Income support)	1	1 May 94	1 Jul 97	No
7 (External training)	1	1 Mar 97	. .	. .
8 (Job Offer)	1	1 Feb 95	. .	. .
8 (Job Offer)	2	1 Dec 95	. .	. .

. . not applicable

Information stored at the episode level is summarised on the Expanded CURF at the person level using summary variables. For example from the information given in the above table, summary variables such as Total duration of working (764 days), Number of looking episodes (1 episode) and Total duration of looking for work (641 days) have been derived.

SETTING OF SIGNIFICANT EPISODES

Although basic information was collected for each episode of labour market activity, detailed information was not collected for a very small number of episodes. While it would have been desirable to collect full details about all episodes, many constraints prevented this. These included the interview situation, and the ability of respondents to recall with accuracy the detail of previous events, particularly for those respondents with numerous episodes.

After all labour market activity episodes had been identified for a respondent in a particular wave, these were ranked according to length of duration (longest to shortest), and the presence of the episode at the end of the previous wave. The highest ranked of these episodes were considered 'significant' and more detailed information was sought about these episodes from the respondent. A maximum quota of up to

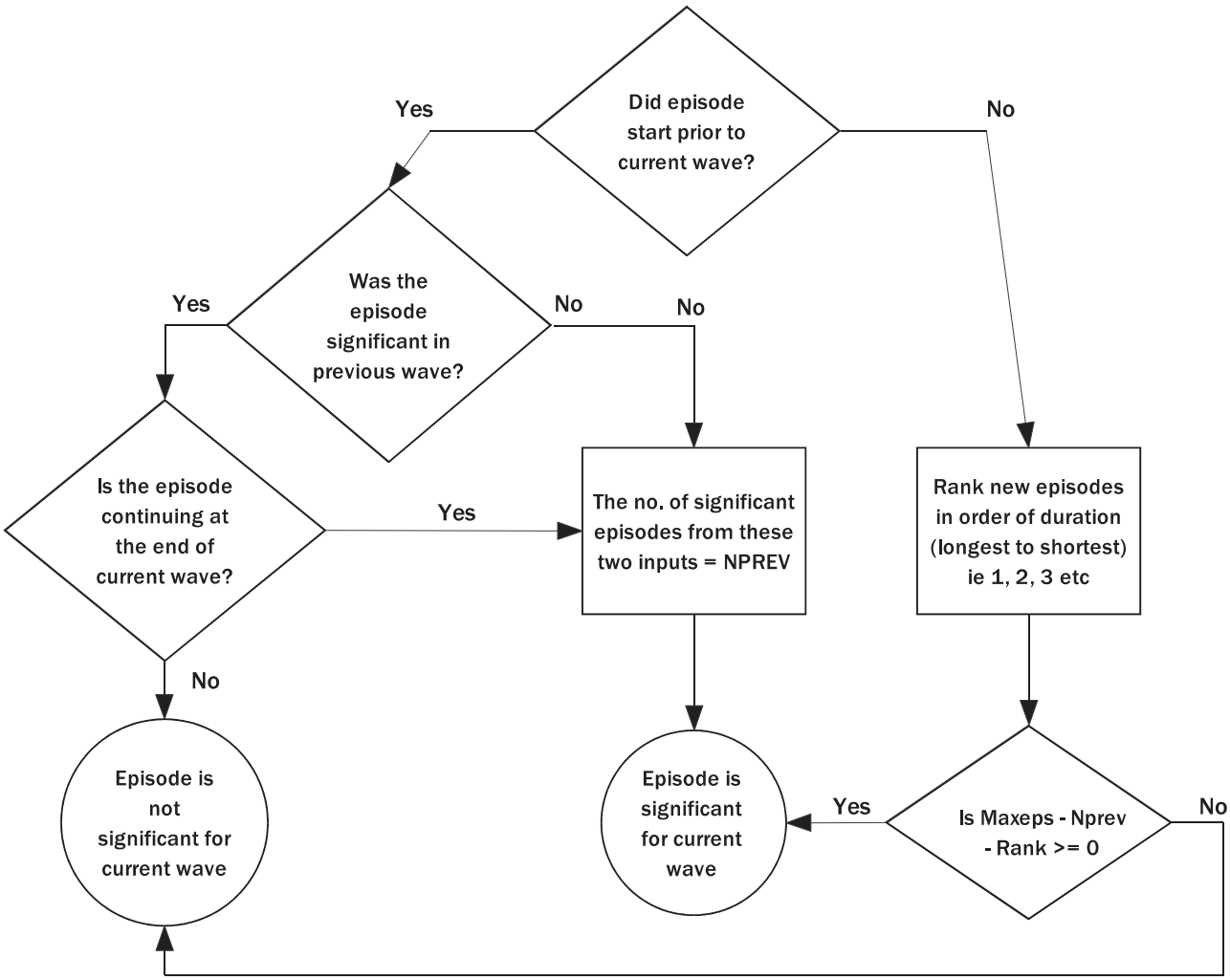
- 6 working episodes;
- 3 looking for work episodes; and
- 3 neither working nor looking for work episodes

could be considered significant. Any remaining episodes in excess of the quota were deemed to be 'insignificant' and more detailed information was not sought about these.

If an episode was significant in the current wave and significant in the previous wave then data from the previous wave was copied to the current wave.

The following diagram (4.7) shows under what conditions an episode is significant for a particular wave.

4.7 Conditions for which an episode is significant for a particular wave



SETTING OF SIGNIFICANT EPISODES *continued*

Where:

- MAXEPS = Maximum number of significant episodes for current wave.
  - 6 for working episodes.
  - 3 for looking and absent episodes.
- NPREV = Number of episodes that started in a previous wave and that are significant in the current wave.

SETTING OF SIGNIFICANT EPISODES *continued*

Table 4.8 shows some examples for setting significant episodes using wave 3 as the current wave.

- Episode 1 is significant for wave 3 as it was significant in wave 2 and is continuing at the end of wave 3.
- Episode 2 is not significant for wave 3 as it was significant in wave 2 and was completed before the end of wave 3. Wave 3 dynamic data items are copied across from wave 2.
- Episodes 3 and 4 are significant for wave 3 as they were not significant in wave 2, but continued into wave 3. Wave 3 dynamic data items are collected, wave 2 dynamic data items remain uncollected.
- Episodes 7, 8 and 9 are significant for wave 3 as:
  - there are a maximum of 6 significant working episodes;
  - three significant working episodes have already been allocated; and
  - of the episodes starting in wave 3, they are the three longest.

4.8 EXAMPLE FOR SETTING SIGNIFICANT EPISODES(a)

Episode Number	*Episode Type*	*Episode started in*	*Is episode continuing at end of wave 3?*	*Was episode significant in wave 2?*	*Episode duration (days)*	*Is episode significant for wave 3?*
1	Working	Wave 1	Yes	Yes	820	YES
2	Working	Wave 2	No	Yes	20	NO
3	Working	Wave 2	Yes	No	420	YES
4	Working	Wave 2	No	No	20	YES
5	Working	Wave 2	Yes	. .	20	NO
6	Working	Wave 3	No	. .	30	NO
7	Working	Wave 3	Yes	. .	40	YES
8	Working	Wave 3	No	. .	50	YES
9	Working	Wave 3	Yes	. .	60	YES

. . not applicable

(a) Where current wave is set at wave 3.

CHAPTER 5 – CONDITIONS OF RELEASE

CONDITIONS OF RELEASE

The SEUP Expanded CURF is released in accordance with a Ministerial Determination (Clause 7, Statutory Rules 1983, No. 19) in pursuance of section 13 of the *Census and Statistics Act 1905*. As required by the Determination, the Expanded CURF has been designed so that the information on the file is not likely to enable the identification of the particular person or organisation to which the data relate.

All organisations and individuals within organisations who request access to a CURF will, prior to being granted access to the CURF, be required to sign an Undertaking to abide by the legislative restrictions on use. The Australian Statistician's approval is required for each release of the CURF. People who buy or access the SEUP CURF are required to give an undertaking which includes, among other conditions, that in using the data they will:

- use the information only for the statistical purposes specified in the Schedule to the Undertaking;
- not attempt to identify particular people or organisations;
- not disclose, either directly or indirectly, the information to any other person or organisation other than members of this organisation who have been approved by the ABS to have individual access to the information;
- not attempt to match, with or without using identifiers, the information with any other list of people or organisations;
- comply with any other direction or requirement specified in the *Responsible Access to ABS CURFs Training Manual*; and
- not attempt to access the information after the term of their authorisation expires, or after their authorisation is rescinded by the organisation which provided it, or after they cease to be a member of that organisation.

Use of the data for statistical purposes means use by people who have signed the Undertaking to produce information of a statistical nature. Examples of statistical purposes are:

- manipulation of the data to produce means, correlations or other descriptive or summary measures;
- estimation of population characteristics;
- use of data as input to mathematical models or for other types of analysis (e.g factor analysis); and
- providing graphical or pictorial representations of the characteristics of the population or subsets of the population.

All CURF users are required to read and abide by the *Responsible Access to ABS CURFs Training Manual* available on the ABS web site <http://www.abs.gov.au> (Services We Provide – CURFs). Use of the data for unauthorised purposes may render the purchaser liable to severe penalties. Advice on the propriety of any particular intended use of the data is available from <curf.management@abs.gov.au> or telephone Canberra (02) 6252 5731.

CONDITIONS OF SALE

All ABS products and services are provided under conditions of sale. Any queries relating to these conditions of sale should be referred to <curf.management@abs.gov.au>.

PRICE

The price to access either the Basic or Expanded SEUP CURF, as at June 2005, is $8,000 (including GST). To access both the Basic and Expanded CURFs, the price is $8,420 (including GST). Clients who already have access to the SEUP Basic CURF (Australians' Employment and Unemployment Patterns, 1994–97, Basic) on CD-ROM can access the SEUP Expanded CURF (Australians' Employment and Unemployment Patterns, (1994–97), Expanded) via the RADL for $420 following submission of an Application and Undertaking for the SEUP Expanded CURF.

HOW TO ORDER

All clients wishing to access the SEUP Expanded CURF should refer to the ABS website, <http://www.abs.gov.au> (Services We Provide – CURFs) and read the *Responsible Access to ABS CURFs Training Manual*, and other relevant information, before downloading the Application and Undertaking to apply for access.

Australian universities

The SEUP CURF can be accessed by universities participating in the ABS/Australian Vice-Chancellors' Committee (AVCC) CURF agreement for research and teaching purposes. For more information, university clients should refer to the ABS web site <http://www.abs.gov.au> (Education Resources – Universities).

APPENDIX 1 STANDARD ERRORS OF ESTIMATES

STANDARD ERRORS OF ESTIMATES

The three following tables provide standard errors for person estimates and episodes of Labour Market activity.

The standard error tables are designed to provide an average standard error applicable to most SEUP person and labour market activity estimates, however the standard errors are not exactly equal for different estimates, and may be quite inaccurate for some unusual estimates.

The size of the standard error increases with the level of the estimate, so that the larger the estimate, the larger the standard error. However, it should be noted that the larger the sample estimate the smaller the standard error will be in percentage terms (that is the relative standard error). Thus, larger estimates will be relatively more reliable than smaller estimates.

As the standard error tables show, the smaller the estimate, the higher the relative standard error. Very small estimates are subject to such high standard errors (relative to the size of the estimate) as to detract from their value for most reasonable uses. In SEUP publications only estimates with relative standard errors of 25% or less, and percentages based on such estimates, are considered sufficiently reliable for most purposes.

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. Standard errors of other estimates and other movements may be determined by using information in *Information Paper: Labour Force Survey Standard Errors* (cat. no. 6298.0) which is available free on the ABS web site <http://www.abs.gov.au> (Key Products – Papers & Articles).

An example of the calculation and use of standard errors is as follows.

Consider an estimate of 200,000 Jobseekers who looked for work only during the wave 2 reference period. By referring to the first table, in the row for an estimate of 200,000 Jobseekers, a standard error of 4,950 is obtained. Therefore, there are about two chances in three that the true value (the number that would have been obtained if the whole population had been included in the survey) is within the range 195,050 (200,000 - 4,950) to 204,950 (200,000 + 4,950). There are about nineteen chances in twenty that the true value is within the range 190,100 to 209,900.

Proportions and percentages (e.g. proportion of Jobseekers who looked for work during the reference period) formed from the ratio of two estimates are also subject to sampling error. The size of the error depends on the accuracy of both the numerator and denominator. The formula for the relative standard error (RSE) of a proportion or percentage is given below:

$$RSE(x/y) = \sqrt{RSE(x)^2 - RSE(y)^2}$$

The following tables contain standard error estimates for:

- Persons, waves 1–3;
- Labour market activities, wave 2; and
- Labour market activities, wave 3.

STANDARD ERRORS OF ESTIMATES *continued*

STANDARD ERRORS, Person estimates: **Waves 1–3**

	JOBSEEKER			POPULATION REFERENCE GROUP		
	Wave 1	*Wave 2*	*Wave 3*	*Wave 1*	*Wave 2*	*Wave 3*
Estimate	no.	no.	no.	no.	no.	no.
100	170	190	200	90	110	100
200	230	250	270	180	220	210
300	270	300	330	280	330	320
500	340	370	400	450	540	540
700	400	430	470	620	730	740
1,000	460	510	540	860	1 000	1 030
1,500	550	600	650	1 230	1 410	1 480
2,000	630	680	750	1 580	1 790	1 900
2,500	700	750	800	1 900	2 140	2 300
3,000	750	810	850	2 200	2 470	2 650
3,500	800	870	950	2 500	2 780	3 000
4,000	850	920	1 000	2 750	3 070	3 350
5,000	950	1 000	1 100	3 300	3 650	3 950
10,000	1 250	1 350	1 450	5 450	5 900	6 600
20,000	1 700	1 850	1 950	8 650	9 300	10 450
30,000	2 050	2 200	2 300	11 150	11 900	13 450
40,000	2 350	2 500	2 600	13 250	14 050	15 900
50,000	2 550	2 750	2 900	15 050	15 900	18 050
100,000	3 500	3 700	3 900	21 750	22 800	25 900
200,000	4 700	4 950	5 200	30 250	31 500	35 600
300,000	5 650	5 900	6 200	35 950	37 350	42 000
500,000	7 050	7 400	7 700	43 800	45 500	50 650
1,000,000	. .	. .	. .	55 250	57 450	62 900
2,000,000	. .	. .	. .	66 950	69 750	74 750

. . not applicable

STANDARD ERRORS OF ESTIMATES *continued*

STANDARD ERRORS, Labour market activity estimates: **Wave 2**

	JOBSEEKERS			POPULATION REFERENCE GROUP		
	Looking	*Working*	*Absent*	*Looking*	*Working*	*Absent*
Estimate	no.	no.	no.	no.	no.	no.
100	220	210	195	100	90	100
200	310	290	275	200	190	205
300	370	360	335	300	290	310
400	430	420	385	400	390	410
500	480	470	430	490	490	505
600	520	510	470	580	590	600
700	560	560	505	670	680	690
800	590	600	540	760	780	780
900	630	630	570	850	870	870
1,000	660	670	605	930	960	955
1,500	800	820	735	1 340	1 400	1 365
2,000	900	950	850	1 700	1 800	1 750
2,500	1 000	1 050	940	2 050	2 200	2 110
3,000	1 100	1 150	1 030	2 400	2 600	2 450
3,500	1 200	1 250	1 110	2 750	2 950	2 780
4,000	1 250	1 350	1 190	3 050	3 300	3 090
4,500	1 350	1 450	1 260	3 350	3 650	3 400
5,000	1 400	1 500	1 330	3 650	3 950	3 690
6,000	1 550	1 650	1 450	4 200	4 600	4 260
8,000	1 750	1 950	1 670	5 200	5 800	5 300
10,000	1 950	2 150	1 870	6 150	6 850	6 260
20,000	2 700	3 100	2 620	10 100	11 400	10 260
30,000	3 250	3 800	3 200	13 250	15 050	13 500
40,000	3 750	4 400	3 690	15 950	18 150	16 280
50,000	4 150	4 900	4 110	18 350	20 900	18 760
100,000	5 750	7 000	5 780	27 750	31 700	28 500
200,000	7 950	10 000	8 120	40 600	46 150	41 890
300,000	9 600	12 300	9 910	49 850	56 450	51 680
400,000	11 000	14 250	11 410	57 300	64 550	59 580
500,000	12 200	15 950	12 730	63 550	71 350	66 260
1,000,000	16 900	22 750	25 150	85 750	94 850	90 220
2,000,000	23 400	32 400	17 890	111 800	121 150	118 850

STANDARD ERRORS OF ESTIMATES *continued*

STANDARD ERRORS, Labour market activity estimates: **Wave 3**

	JOBSEEKERS			POPULATION REFERENCE GROUP		
	Looking	*Working*	*Absent*	*Looking*	*Working*	*Absent*
Estimate	no.	no.	no.	no.	no.	no.
100	260	280	205	100	90	105
200	360	400	290	200	190	215
300	430	480	355	310	300	320
400	490	550	410	410	400	420
500	550	620	460	500	510	520
600	600	680	505	600	610	620
700	640	730	545	690	710	715
800	680	780	580	780	800	805
900	720	820	615	870	900	900
1,000	760	870	650	960	990	985
1,500	920	1 060	795	1 370	1 450	1 410
2,000	1 050	1 200	920	1 750	1 900	1 810
2,500	1 150	1 350	1 030	2 100	2 300	2 180
3,000	1 250	1 500	1 130	2 450	2 700	2 530
3,500	1 350	1 600	1 220	2 800	3 050	2 870
4,000	1 450	1 700	1 300	3 100	3 400	3 200
4,500	1 550	1 800	1 380	3 400	3 750	3 510
5,000	1 600	1 900	1 450	3 700	4 100	3 820
6,000	1 750	2 100	1 590	4 300	4 750	4 400
8,000	2 000	2 400	1 840	5 350	6 000	5 480
10,000	2 200	2 650	2 050	6 300	7 100	6 470
20,000	3 050	3 750	2 900	10 350	11 750	10 610
30,000	3 700	4 550	3 550	13 550	15 550	13 960
40,000	4 250	5 250	4 100	16 350	18 750	16 840
50,000	4 700	5 850	4 580	18 800	21 650	19 400
100,000	6 500	8 200	6 470	28 450	32 750	29 470
200,000	9 000	11 500	9 150	41 550	47 700	43 320
300,000	10 900	14 000	11 200	51 050	58 350	53 440
400,000	12 450	16 100	12 930	58 650	66 750	61 600
500,000	13 800	17 950	14 450	65 100	73 800	68 510
1,000,000	19 100	25 200	20 420	87 800	98 050	93 280
2,000,000	26 400	35 350	28 850	114 450	125 300	122 880

ANALYSIS CASE STUDIES

The purpose of this section is to illustrate, via simple case studies, the type of data that is available on the Expanded CURF, the data's longitudinal nature and the way in which the data can be manipulated and used.

The case studies involve analysing data at both the person and episode level and are presented in increasing order of complexity.

The outputs of the case studies have been weighted based on the standard Jobseeker weight JSWGT3. This weight was selected as the weighted frequencies provided in the documentation file (DATA ITEM AND FREQUENCY LISTING.TXT) are based on Jobseeker weights. The program output can then be compared against the weighted frequencies in the documentation file to provide a convenient means of validating the program code. Once the program code has been validated using this method the program can then be re-run using whatever weights are appropriate.

CASE STUDY 1

Age distribution of all Jobseekers compared to that of Jobseekers who were in employment at the end of wave 3.

Variables used

Person record variables:

- AGES (Age at May 1995)
- LFSEUPS3 (Labour Force Status at end of wave 3)
- JSWGT3 (Jobseeker weight for wave 3)

Method

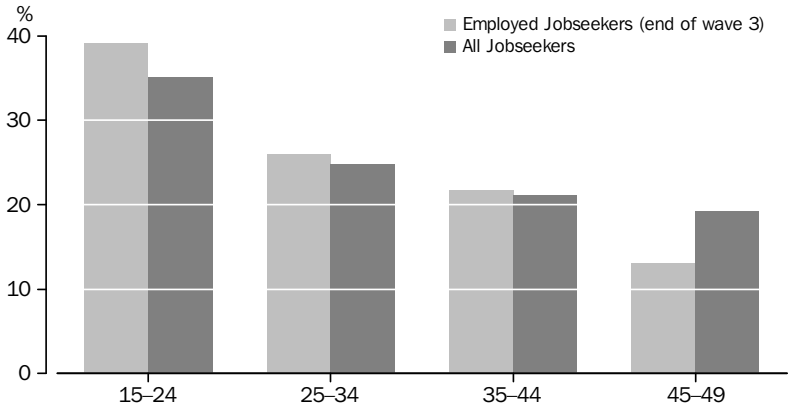
1. Identify Jobseekers by selecting records with JSWGT3 > 0.

2. Use your preferred analysis tool to get a distribution of AGES for Jobseekers.

3. Identify Employed Jobseekers by selecting records with JSWGT3 > 0 and LFSEUPS3 = 1 or 2 (Employed full time or part time).

4. Use your preferred analysis tool to get a distribution of AGES for selected records.

1.A AGE DISTRIBUTION OF JOBSEEKERS



Validating the output

The output shown in graph 1.A is based on the figures shown in table 1.B.

1.B AGE DISTRIBUTION OF JOBSEEKERS

Age (years)	Jobseekers who found employment no.	All Jobseekers no.	Jobseekers who found employment %	All Jobseekers %
15–24	178 013	306 222	39.2	35.0
25–34	118 213	216 737	26.0	24.8
35–44	98 784	184 544	21.7	21.1
45–49	59 342	167 594	13.1	19.2
Total	454 352	875 097	100.0	100.0

Weighted frequencies for Jobseekers can be validated by looking at the weighted age distribution in the documentation file DATA ITEM AND FREQUENCY LISTING.TXT.

CASE STUDY 2

Change in main income source from wave 1 to wave 3

Variables used

Person record variables:

- MSRCINC1 (Main source of income in wave 1)
- MSRCINC3 (Main source of income in wave 3)
- JSWGT3 (Jobseeker group weights for wave 3)

Method

This case study estimates the number of Jobseekers whose main source of income changed from income support in wave 1 to some other form of income. By way of comparison the case study also estimates the number of Jobseekers whose main source of income support in wave 1 changed from wages and salaries to another form of income.

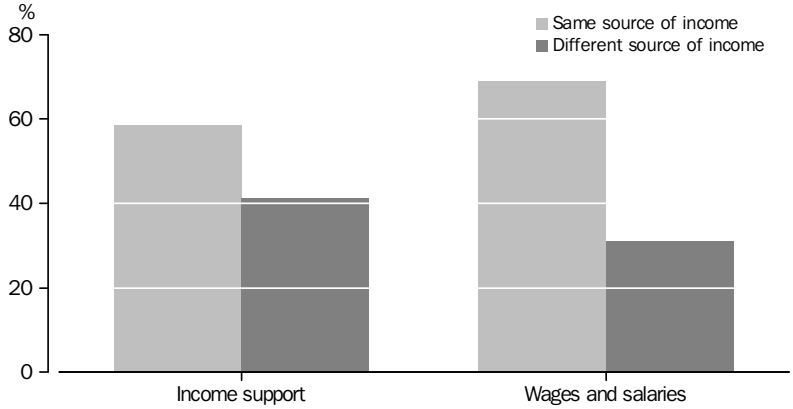
1. Identify Jobseekers whose main source of income in wave 1 was income support from the government, by selecting records with JSWGT3 > 0 and MSRCINC1 = 5.

2. Use your preferred analysis tool to get a distribution of MSRCINC3 for these records. Records with MSRCINC3 not equal to 5 have changed their main income source.

3. Identify Jobseekers whose main source of income in wave 1 was wages and salaries from employers by selecting records with JSWGT2 > 0 and MSRCINC1 = 1.

4. Use your preferred analysis tool to get a distribution of MSRCINC3 for these records. Records with MSRCINC3 not equal to 1 have changed their main income source between September 1995 and September 1997.

2.A CHANGES IN MAIN INCOME SOURCE, From wave 1 to wave 3



Method continued

2.B CHANGES IN MAIN SOURCE OF INCOME

Main source of income in wave 1	*Same source of income in Wave 3* no.	*Different source of income in Wave 3* no.	*Total* no.	*Proportion with same source of income in Wave 3* %	*Proportion with different source of income in Wave 3* %
Income support	337 936	238 037	575 973	58.7	41.3
Wages and salaries	135 818	60 902	196 719	69.0	31.0

The total number of people on income support in the above table can be validated against the documentation file. It should equal the number of people on income support for data item MSRCINC3, weighted on JSWGT3. The number of wage and salary earners can be found in the same way.

CASE STUDY 3

Main activities of Jobseekers who were absent from the labour market during the survey

Variables used

Episode record variables:

- MNACTYN1 (Main Activity in wave 1)
- MNACTYN2 (Main Activity in wave 2)
- MNACTYN3 (Main Activity in wave 3)
- EPSTAT (Episode Status)
- JSWGTQ3 (Jobseeker weights for wave 3)

Method

1. Select episodes of absence from the labour market for Jobseekers, i.e. records with EPSTAT =3 and JSWGTQ3 > 0. Some of these episodes will span all three survey waves, others will span only one or two waves. Records that span all survey waves will have a response for MNACTYN1, MNACTYN2 and MNACTYN3. Records that are only in one or two waves will have some of these variables set to zero.

2. To select the appropriate variable for analysis implement the following logic:

IF MNACTYN3 > 0 THEN use MNACTYN3

ELSE IF MNACTYN2 > 0 THEN use MNACTYN2

ELSE use MNACTYN1

This will select the wave 3 variable for episodes that are exclusively in the wave 3 reference period or which span wave 1 through wave 3. The wave 2 variable will only be selected for episodes that only span waves 1 and 2 or that fall exclusively within the wave 2 reference period. The wave 1 variable will only be selected for episodes that fall exclusively within the wave 1 reference period.

3. Use your preferred analysis tool to get the following distribution for main activity.

Method continued

3.A DISTRIBUTION OF MAIN ACTIVITIES, Jobseekers

Main activity	*Number of episodes for Jobseekers* no.	*Proportion* %
Retired or voluntarily inactive	29 866	4.0
Home duties/childcare	220 705	29.7
Study/went to school/TAFE/university	209 822	28.2
Own illness/injury	89 956	12.1
Own disability/handicap	7 824	1.1
Looking after ill/disabled person	17 696	2.4
Travel/moving house/holiday	108 431	14.6
Working in unpaid voluntary job	9 184	1.2
Other activity	49 860	6.7

Validating the output

The number of episodes for each main activity from wave 1 to wave 3 should be greater than the number of episodes for each individual wave. However as some episodes are in all three waves and other episodes span two waves, the number of episodes across wave 1 to wave 3 should be less than the sum for all waves.

If the main activity is a holiday, the duration of this episode would tend to be less than other activities. Thus there would be less likelihood of an episode spanning more than one wave, and therefore the sum of wave 1 to wave 3 episodes would be closer to the number of episodes across wave 1 to wave 3.

The following table shows the number of episodes across wave 1 to wave 3 with a response to main activity. It also shows the number of episodes with a response to main activity for each wave. The ratio is obtained by summing the number of episodes for each wave and dividing it by the number of episodes across wave 1 to wave 3.

3.B MAIN ACTIVITIES OF JOBSEEKERS, Number of episodes: **Wave 1–wave 3**

	NUMBER OF EPISODES			
	Jobseekers	*Wave 1*	*Wave 2*	*Wave 3*
Main activity	no.	no.	no.	no.
Retired or voluntarily inactive	29 866	13 367	9 322	17 694
Home duties/childcare	220 705	127 524	91 566	102 061
Study/went to school/TAFE/university	209 822	147 986	74 646	47 238
Own illness/injury	89 956	38 482	48 367	49 391
Own disability/handicap	7 824	2 218	4 784	6 813
Looking after ill/disabled person	17 696	8 579	10 408	9 779
Travel/moving house/holiday	108 431	63 030	27 092	24 559
Working in unpaid voluntary job	9 184	4 089	2 498	4 849
Other activity	49 860	25 200	24 824	12 599

CASE STUDY 4

Jobseeker Labour Market Transitions, September 1994 to September 1997

Variables used

Person record variables:

- LMABEG1 (Labour Market Activity at beginning of wave 1)
- LMAEND3 (Labour Market Activity at end of wave 3)
- JSWGT3 (Jobseeker weight for wave 3).

Method

This case study categorises Jobseekers as either working, looking for work or absent from the labour market at September 1994. The predominant Labour Market Activity of people from each of these activities at September 1997 is then determined, indicating whether there has been any transition between labour market activities.

1. Extract records from the Person file with JSWGT3 > 0.

2. Use your preferred analysis tool to obtain a distribution of LMABEG1 by LMAEND3 weighted on JSWGT3.

4.A JOBSEEKER LABOUR MARKET TRANSITIONS

	AT END OF WAVE 3				
	Working	*Looking*	*Absent*	*Working & Looking*	*Sum of wave 1 responses*
	no.	no.	no.	no.	no.
At start of wave 1					
Working	131 762	41 462	28 333	21 652	223 209
Looking	137 869	170 372	86 043	40 729	435 012
Absent	77 655	40 594	45 014	16 651	179 914
Working & Looking	18 829	7 070	3 355	7 708	36 962
Sum of wave 3 responses	366 114	259 497	162 745	86 740	875 097

Validating the output

The sum of wave 3 responses (i.e. the bottom row of the above table) should match the distribution of LMAEND3 in the documentation file, weighted on JSWGT3. Likewise the sum of wave 1 responses (the last column in the table) should match the distribution of LMABEG1.

APPENDIX 3 SPECIAL CODES FOR CONTINUOUS DATA ITEMS

CONTINUOUS DATA ITEMS, Special codes: **Persons and Episodes**

Field name and label	*Response*	*Code*
	PERSONS	
Duration of employment of last full-time job in months		
EMPDURFT	Under 1 month	9997
Duration of employment of last part-time job in months		
EMPDURPT	Not available	9998
EMPDURPT	Under 1 month	9997
Finish date of last full-time job		
FINDATFT	Not applicable	999999
Finish date of last part-time job		
FINDATPT	Not applicable	999999
Number of years in paid work, since first left full-time education		
YPWLFTES	None	96
YPWLFTES	Not applicable - still studying	0
YPWLFTES	Less than 1 year	95
Number of years spent looking for work, since first left full-time education		
YLWLFTES	Less than 1 year	95
YLWLFTES	None	96
YLWLFTES	Not applicable - still studying	0
Number of years spent neither working nor looking for work, since first left full-time education		
YNWLFTES	Not applicable - still studying	0
YNWLFTES	None	96
YNWLFTES	Less than 1 year	95
Start date of last full-time job		
STADATFT	Not applicable	999999
Start date of last part-time job		
STADATPT	Not applicable	999999
Time since finished last full-time job		
TSFLFT	Under 1 month	9997
Time since finished last part-time job		
TSFLPT	Under 1 month	9997
TSFLPT	Not available	9998
Total duration of case management in reference period		
DURCAS1	No consent given	999
DURCAS2	No consent given	999
DURCAS1	Consent given but no records	997
DURCAS2	Consent given but no records	997
Total duration of CES registration in reference period		
DURCES1	No consent given	999
DURCES2	No consent given	999
DURCES1	Consent given but no records	997
DURCES2	Consent given but no records	997

CONTINUOUS DATA ITEMS, Special codes: **Persons and Episodes** *continued*

Field name and label	*Response*	*Code*
PERSONS *cont.*		
Total duration of LMP in reference period		
DURLMP1	No consent given	999
DURLMP2	No consent given	999
DURLMP1	Consent given but no records	997
DURLMP2	Consent given but no records	997
Total time spent on external training course in hours		
TIMEXT59	Not stated	9999
Total time spent on in-house training course in hours		
TIMEIN59	Not stated	9999
Total time spent on (4 longest) external training course in hours		
TIMEXTS1	Not stated	9999
TIMEXTS2	Not stated	9999
Total time spent on (4 longest) in-house training course in hours		
TIMEINT1	Not stated	9999
TIMEINT2	Not stated	9999
Years since first left full-time education		
YLFTES	Not stated	99
YLFTES	None	96
YLFTES	Less than 1 year	95
EPISODES		
Duration of external training in hours		
EPTHLMA7	Not stated	9999
Duration of in-house training in hours		
EPTHLMA6	Not stated	9999
Episode beginning date		
BEGDATE	Not stated	999999
Episode end date		
ENDDATE	Not applicable	999999
When offer of employment was made		
JOBOFFMD	Not stated	9999

GLOSSARY

Absent from the labour market — Neither working nor looking for work.

Business size — The total number of people who work for an employer at all locations.

Casual job — A job in which the employed person is not entitled to either paid annual leave or paid sick leave.

Episode — A period of time during which a particular activity is undertaken (such as working, looking for work or absence from the labour market). An episode can occur wholly within a reference period or can span more than one reference period.

Full-time work — Work in which the employed person usually works 35 hours or more a week in all jobs.

Income — The income received within a given time period before tax or any other deductions are made. It includes wages or salary, pensions, benefits and allowances, business, investment and property income.

Industry — All occurrences of industry refer to Industry Division as defined by the *Australian and New Zealand Standard Industrial Classification* (ANZSIC) (cat. no. 1292.0).

Jobseekers — The Jobseeker component of the SEUP sample consists of people aged 15–59 years who met the following criteria at the time of recruitment to the survey panel (24 April 1995 to 7 July 1995):

- unemployed people;
- people not in the labour force who were discouraged Jobseekers;
- people not in the labour force who were attending an educational institution;
- people not in the labour force who wanted to work but were not available to start work; or
- underemployed workers.

Labour market activities — Periods of working, looking for work, and absence from the labour market.

Main English-speaking countries — Comprises the United Kingdom, Ireland, Canada, South Africa, the United States of America and New Zealand.

Minimal work experience — Less than six months work between May 1995 and September 1997.

Moderate work experience — At least six months work between May 1995 and September 1997, with an average job duration of between six and 12 months.

Occupation — All occurrences of occupation refer to Major Group as defined by the *Australian Standard Classification of Occupations, Second Edition* (cat. no. 1220.0).

Part-time work — Work in which the employed person usually works less than 35 hours a week in all jobs.

Permanent job — A job in which the employed person is entitled to paid annual leave or paid sick leave.

Shorter work experiences — At least six months work between May 1995 and September 1997 (in total, for all jobs), with an average job duration of less than six months.

Stable job — Job which lasted for six months or more and the jobholder was not concurrently looking for other work.

Sustained work experiences — Jobseekers who worked for at least 12 months between May 1995 and September 1997, with an average job duration of at least 12 months.

Unstable job — Job which lasted for less than six months or the jobholder was concurrently looking for other work.

Wage and salary earner — A person who works for a public or private employer and receives remuneration in wages or salary. Excludes people in their own business, either with or without employees, if that business was incorporated.

Wave Describes the reference period for each data collection. The reference periods are as follows:

Wave 1 – 5 September 1994 to 3 September 1995;

Wave 2 – 4 September 1995 to 1 September 1996; and

Wave 3 – 2 September 1996 to 31 August 1997.

Other definitions For definitions of labour force and demographic classifications used in this publication, see *Australian Labour Market Statistics* (cat. no. 6105.0), and *Labour Statistics: Concepts, Sources and Methods* (cat. no. 6102.0.55.001) which is available on the ABS web site. <http://www.abs.gov.au> (Methods, Classifications, Concepts & Standards - ABS concepts, classifications and statistical standards).

FOR MORE INFORMATION . . .

INTERNET	**www.abs.gov.au** the ABS web site is the best place to start for access to summary data from our latest publications, information about the ABS, advice about upcoming releases, our catalogue, and Australia Now—a statistical profile.
LIBRARY	A range of ABS publications is available from public and tertiary libraries Australia-wide. Contact your nearest library to determine whether it has the ABS statistics you require, or visit our web site for a list of libraries.
CPI INFOLINE	For current and historical Consumer Price Index data, call 1902 981 074 (call cost 77c per minute).
DIAL-A-STATISTIC	This service now provides only current Consumer Price Index statistics call 1900 986 400 (call cost 77c per minute).

INFORMATION SERVICE

Data already published that can be provided within five minutes will be free of charge. Our information consultants can also help you to access the full range of ABS information—ABS user pays services can be tailored to your needs, time frame and budget. Publications may be purchased. 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

WHY NOT SUBSCRIBE?

ABS subscription services provide regular, convenient and prompt deliveries of selected ABS publications and products as they are released. Email delivery of monthly and quarterly publications is available.

PHONE	1300 366 323
EMAIL	subscriptions@abs.gov.au
FAX	(03) 9615 7848
POST	Subscription Services, ABS, GPO Box 2796Y, Melbourne Vic 3001

2000001492635

RRP $10.00