Foreword

Public sector agencies rely on high-quality and up-to-date data to provide innovative, evidence-based advice to government. Australia’s longitudinal studies can make a particularly valuable contribution to this because their data shows what matters and works.

Recognising this potential, over the past decade, Australian Government Department of Families, Housing, Community Services and Indigenous Affairs and other Commonwealth agencies have established several large-scale Australian longitudinal studies.

Data from these studies can show the factors that make a difference to people’s outcomes over time. The breadth of domains that many longitudinal datasets cover—including education, health, early development, multiple and deep disadvantage, housing, work, family and community—provides a unique, cross-portfolio perspective for understanding the often complex and interacting issues facing policy-makers.

This Guide to Australian longitudinal studies highlights the importance of such Australian longitudinal data as a national resource. The value of this resource continues to be enriched as new waves of data are added. This guide provides an introduction to some of the possibilities that longitudinal research opens up and guidance for starting to use the data.

By building awareness and capacity, we can leverage the growing potential of longitudinal data to ensure policy advice best meets the short- and long-term challenges ahead.

Serena Wilson
Deputy Secretary

Department of Families, Housing, Community Services and Indigenous Affairs
1 Introduction

Longitudinal datasets offer a powerful resource for decision-makers. Despite the growth in the availability and analysis of longitudinal data, its potential to support policy has not been fully realised.

This Guide to Australian longitudinal studies aims to help redress this by showcasing why longitudinal datasets are so valuable, what datasets are available and how to use them to better inform policy.

The guide explains the unique strengths and opportunities of longitudinal data, including why longitudinal data allows decision-makers to address critical questions that more widely available cross-sectional datasets cannot. The guide also provides practical examples of how longitudinal data can improve decision-making, summarises the features of some of Australia’s major longitudinal studies and provides links to a range of further resources for helping agencies to understand and use the data.

This guide was developed by the Research and Analysis Branch within Australian Government Department of Families, Housing, Community Services and Indigenous Affairs with significant input from the Australian Bureau of Statistics (ABS). Other government agencies and research institutes were also consulted and contributed to the guide, including the Australian Institute of Family Studies (AIFS), the Australian Institute of Health and Welfare (AIHW), the Department of Education, Employment and Workplace Relations (DEEWR), the Department of Health and Ageing (DoHA), the Department of Immigration and Citizenship (DIAC), the Melbourne Institute, and the Telethon Institute for Child Health Research.
2 What is longitudinal data?

Longitudinal data is data collected from the same individuals or households at several points in time. The data allows differences between and within individuals to be tracked over time.

Longitudinal data differs from the cross-sectional data that most people are familiar with, such as the census. A cross-sectional study provides a snapshot of the numbers of people experiencing an event or circumstance, but only at a single point in time. A series of repeated cross-sections, or time-series, can show broad population trends over time. However, unlike longitudinal data, these studies cannot show changes experienced by an individual over time.

If cross-sectional datasets are ‘snapshots’ of the population, then longitudinal datasets provide ‘moving pictures’ showing the individual journeys different people take and how their lives unfold through time.

Better evidence demands both better data and better methodologies. Relative to the USA and UK, Australia has made little use of longitudinal data, which can be of crucial importance in assessing the relative significance of different influences on observed outcomes.

3 How does longitudinal data support good decision-making?

Analysing cross-sectional data is an important first step in identifying and describing policy problems. For example, the census provides the best estimates of the numbers of people in the population with particular characteristics, such as education levels or housing status.

Longitudinal data allows decision-makers to go beyond describing the extent of a policy problem to develop some understanding of how and why problems occur and what is likely to help. It also enables identification of the consequences of problems. Because longitudinal data tracks individual pathways, it can show how different people respond to opportunities and setbacks, and how their responses and outcomes change in the short and long term. This provides policy-makers a breadth and depth of perspective—showing how circumstances and policy levers affect outcomes over time.

Uniquely, longitudinal data can help decision-makers:

- understand persistence and change in people’s circumstances
- know what makes a difference to people’s outcomes
- understand how experiences interact over time (both long and short term) and across different life domains (health, education, employment, family, social)
- test ideas and evaluate policies and programs.

Snapshot (cross-sectional) data can lead people to focus on the symptoms of the problem rather than addressing the underlying processes which lead people to have or be denied opportunities. To understand why people’s life chances differ, it is important to look for the events and experiences which create opportunity and those which create barriers, and to use this as a focus for policy action.

4 Understanding persistence and change in people’s circumstances

Experiencing disadvantage, such as poor health or poverty, can be a temporary setback for some, but for others such problems can become entrenched. In each case a very different policy response is required. To measure persistence, cross-sectional surveys could, in theory, ask people to recall their earlier circumstances. However, research shows that these recall measures tend to be biased and unreliable, particularly over longer periods.

Only longitudinal data can reliably show whether circumstances are:

- temporary
- persistent (and if so, for how long)
- recurrent/relapsing (and if so, how often relapse occurs).

Longitudinal data allows researchers to identify factors that influence transition and persistence. Some examples of the kinds of question that longitudinal data can illuminate include:

- what supports the long-term unemployed to get and keep a job?
- what factors influence the course of recurrent mental health problems?
- how do people change their asset portfolios before, during and after retirement and in response to government policies?

Understanding persistence in income poverty

Distinguishing between people facing temporary spells of poverty from those experiencing long-term financial disadvantage is critical for designing and targeting effective social policies and payments. Data from the Household Income and Labour Dynamics in Australia (HILDA) Survey, for example, can show the patterns of persistence in relative income poverty—often defined as a household income of less than 50 per cent of the median. Research using this data shows that around 5 to 6 per cent of Australians typically enter relative income poverty in any given year, a similar proportion exit poverty each year, and a further 7 to 8 per cent are in relative income poverty over any two years. The research showed that of those in relative poverty in 2001, approximately 43 per cent were in poverty 5 years later and 48 per cent were in poverty 8 years later. Without longitudinal data it would not be possible to show these patterns of income poverty persistence over the short and medium term.

5 Understanding what makes a difference to people’s outcomes

Effective policy necessarily targets the causes of problems. However for many complex policy problems it is not easy to determine the size and direction of the different causal influences that really matter.

While randomised control trials, such as the ‘pill-placebo’ studies used in medical research, are the gold standard for demonstrating causal relationships, it is often impractical or unethical to run such studies to assess policy questions.

Longitudinal data often provides the next best available evidence for quantifying the relative strengths of different causal influences on an outcome. Longitudinal data can show the temporal sequence of possible causes and effects and so provide evidence that is consistent with a causal relationship. More advanced ‘fixed effects’ techniques allow researchers to control for otherwise unmeasured, stable individual characteristics—such as temperament, or differences in ability—that cannot be controlled for using cross-sectional data.

Without knowing and quantifying these causal processes, it is not possible to know whether, and to what extent, a policy designed to achieve an outcome (e.g. better mental health of children) by improving some situation (parenting) is likely to work.

Does a ‘jobs first’ approach really put people at greater risk of future unemployment?

A question that is critical for employment policy is whether taking a low-skilled job reduces the risk of future unemployment (the so-called ‘jobs first’ approach) or merely leads to spells of low pay and unemployment (the ‘low-pay, no-pay’ cycle). Researchers examined seven waves of HILDA Survey data and found, after controlling for selected observed and unobserved background characteristics, that low paid employment is generally protective against future spells of unemployment—consistent with the ‘jobs first’ approach to employment policy. Importantly, by using longitudinal data, the researchers were able to examine employment transitions and control for unobservable background characteristics that would have not been possible using cross-sectional data.

6 Explaining how experiences interact over time and across different life domains

Many longitudinal studies collect information across a breadth of life stages and domains, such as early childhood, parenting, education, employment, physical and mental health, finances, housing, social relationships and family. In addition, longitudinal data can show how these domains interact across the life span; from as early as before an individual is born right through to subsequent generations.

By exploiting this breadth and depth of data, longitudinal studies are particularly powerful tools for identifying sensitive windows of development when the return on investment from policy intervention can be very high. The benefits of early intervention and prevention can also spread across portfolio outcomes. For example, longitudinal data has underpinned the now widely understood social and economic benefits for later educational, mental health and employment outcomes of investment in early childhood.

By informing early intervention strategies, longitudinal data is uniquely placed to help address problems that are much harder and more expensive to change later in life.

Post-natal mental health: an issue for both mothers and fathers

Much previous research has examined the causes and consequences of post-natal depression for mothers and their children, but relatively little work has explored these same issues for fathers. More recent research using data from the Longitudinal Study of Australian Children (LSAC) has shown that around 10 per cent of fathers experience clinically significant psychological distress in the post-natal period. This is about one and a half times the rate in the general male population and comparable to the rates of post-natal depression in mothers. LSAC research also shows that, similarly to mothers, fathers’ post-natal depression goes on to negatively affect children’s social and emotional development over time—even until children are aged five years.


7 Testing ideas and evaluating policy

Longitudinal data can help decision-makers test and cost policy ideas and evaluate smaller-scale interventions.

Comprehensive longitudinal data can help quantify the cumulative costs of policy action (or inaction). A number of longitudinal studies include linkages with administrative data (such as Medicare data) that allow policy-makers to establish the costs and potential cost-effectiveness of different policy options. Many longitudinal datasets also include indicators that are frequently used to monitor or compare policies, for example National Assessment Plan Literacy and Numeracy (NAPLAN) data in education, Quality Adjusted Life Years (QALY) and Health Related Quality of Life measures in healthcare and Australian Early Development Index (AEDI) data in early child development.

Longitudinal studies can also serve as a benchmark for evaluating smaller-scale, trial interventions. By matching individuals within a small-scale intervention trial with individuals sharing similar observed characteristics from a longitudinal study (who have not experienced the intervention, i.e. the ‘control group’) it is possible to estimate how effective the intervention is compared to the status quo.

Evaluating the Home Interaction Program for Parents and Youngsters (HIPPY)

The HIPPY program supports parents from disadvantaged backgrounds to read to and guide their pre-school age children’s early learning experiences and so better prepare them for school. In 2011 researchers compared outcomes of families who had completed the program to similar, “matched” children from the LSAC study who had not completed the program.

The evaluation showed a range of positive findings, including:

- The parenting style of HIPPY parents was significantly less angry or hostile compared to similar LSAC parents.
- HIPPY parents did significantly more in-home and out-of-home activities with their child compared to similar LSAC parents.
- HIPPY children’s learning and literacy levels improved.
- The program showed a predicted return to society from improved employment and educational outcomes of at least $2.53 for every $1 spent (assuming a 3% discount rate).

8 Types of longitudinal study

Although all longitudinal studies share the distinguishing feature of linking individual data records collected at different points in time, there is a wide range of designs of longitudinal study, each with different areas of focus.

- Household panel surveys such as the HILDA Survey repeatedly collect data from the same households over many years and typically focus on the dynamics of economic and subjective wellbeing, employment and families.
- Birth cohort studies (e.g. LSAC’s B cohort) collect data from the families of children in their first year of life (or before) and follow them over time to investigate how early-life experiences (childcare, early health, schooling etc.) influence educational, health and employment outcomes in later life.
- Age cohorts are similar to birth cohort studies but respondents are older when data is first collected. For example the LSAC K cohort were aged 4–5 years in the first wave of data collection, and LSAY cohorts are aged 15 years at Wave 1.
- Special population studies focus on the experience of individuals with a particular background, such as recently-arrived migrants to Australia, or women (as in the Australian Women’s Health Study).
- Area studies involve the experiences of individuals from a particular location and are commonly used to investigate physical health or to investigate the experience of local institutions or services (e.g. the Stronger Families in Australia study).
- Other specialised longitudinal studies can track other groups for analysis, for example the Business Longitudinal Survey surveys a cohort of small to medium enterprises over time.
- Longitudinal studies can also be constructed from integrated datasets using administrative and/or survey data.
9 Assessing the quality of longitudinal data for analysis purposes

When considering the use of longitudinal data, the quality and relevance of the data for the intended analysis should be determined. Key quality issues include:

- whether the original sample is representative of a population of interest and thus allows inferences regarding that population
- any loss of sample units (such as people, households or businesses) from the sample at each wave in the longitudinal dataset
- whether the dataset fails to include or under-represents additions to the population since the original participants were selected (for example significant change in migration patterns)
- whether, over time, previous experience with the survey may lead to a change in reporting or behaviour
- ensuring any changes in the external environment (for example changes in government payments) are taken into account before using the data to predict future trajectories.

Supporting documentation for longitudinal datasets should provide background information to determine the appropriate use of the data.
10 Commissioning longitudinal research and training

Many universities and private research providers across Australia have researchers with specialist expertise in conducting longitudinal data analysis.

Commonwealth Accredited Integrating Authorities have expertise in constructing longitudinal integrated datasets for analysis from existing administrative and survey data. For further information see the National Statistical Service Website: www.nss.gov.au/dataintegration.

Several research institutes offer training in longitudinal data analysis techniques, including the Australian Consortium for Social and Political Research Incorporated (ACSPRI; http://www.acspri.org.au/), the Melbourne Institute (http://www.melbourneinstitute.com), and the National Centre for Social and Economic Modelling (NATSEM; http://www.natsem.canberra.edu.au/courses/).

For further information on how to commission longitudinal research, including access to researchers from FaHCSIA’s Social Policy Research and Evaluation Panel (SPREP) and information on longitudinal analysis training providers, please contact the Research and Analysis Branch within FaHCSIA at: research.management@fahcsia.gov.au.
Household, Income and Labour Dynamics in Australia (HILDA) Survey

The Household, Income and Labour Dynamics in Australia (or HILDA) Survey is one of only a small number of well-established, large, nationally representative, household-based panel studies conducted in the world.

Who is involved?
The wave 1 panel consisted of 7682 households and 19,914 individuals. In wave 11, this was topped up with an additional 2153 households and 5451 individuals. The survey commenced in 2001 and, with annual data collection, there is now over 10 years of unit record data available to researchers.

The HILDA Survey was initiated, and is funded, by FaHCSIA. Responsibility for the design and management of the survey rests with the Melbourne Institute of Applied Economic and Social Research (University of Melbourne). The Nielsen Company collected data for waves 1 to 8, and data collection for waves 9 to 16 is being undertaken by Roy Morgan Research.

Content areas
The content of the HILDA Survey instruments focuses on: changes to families, including family formation and fertility intentions; incomes and wealth; employment, including retirement intentions; health; education; and life satisfaction, disability and social capital.

The new-person questionnaire, for people who have not previously responded, has the same content but, in addition, collects information about respondents’ backgrounds, including: country of birth and language, visa status and migration category (if born overseas); educational attainment; employment history; marital history; family background; and parents’ characteristics.

Each wave of the HILDA Survey includes at least one major topic for particular focus, which is repeated every four years. These topics include: household wealth, family formation and fertility, retirement from the workforce, health and education, skills and abilities. In addition, there are a small number of shorter questions included on a rotating basis, including: job related discrimination (included in waves 8 and 10, and next scheduled to be included in wave 14); intentions and plans regarding mobility, education and work (waves 5, 8 and 11 and thereafter to be moved on to a four-year cycle); and non-coresidential family relationships (waves 8 and 12).

Further information
Further information about the study, including data access, the data dictionary and links to publications and reports, is available at: http://www.melbourneinstitute.com/hilda/
Benefits of vocational education for helping people with disability to gain employment

People with disability still experience significant barriers to realising their potential to participate in social activities and employment. For example, research using HILDA data has shown that people with disabilities who are unemployed and have not received vocational education are at high risk of remaining unemployed over the following year (74%), and after three years (50%). However, the research also showed that receiving vocational education causes a large and significant reduction in these risks and that these benefits last for at least three years after training completion. The Productivity Commission Report on Disability Care and Support drew on these longitudinal findings to emphasise the critical importance of vocational education, and other supports, for improving the participation outcomes of people with disability.

Longitudinal Study of Australian Children (LSAC)

Growing up in Australia: the Longitudinal Study of Australian Children (LSAC) is the first comprehensive, nationally representative, Australian dataset on children as they grow up.

Who is involved?
The study follows two cohorts of children, with approximately 5000 children in each. (in wave 1). Parents, child care providers, teachers and the children themselves provide information and the ABS interviewers undertake direct observations and assessments.

For the first wave of the study, in 2004, the infant cohort was aged 3–17 months and the child cohort was aged between 4 and 5 years. Study families are visited for a face-to-face interview every two years.

LSAC is conducted by a partnership comprising FaHCSIA, AIFS and the ABS, with advice provided by a consortium of leading researchers.

Content areas
At each wave, a wide range of age-appropriate developmental outcomes are measured, including children’s health, physical development, emotional wellbeing and intellectual and social development. In addition, data is collected on the context in which development is occurring, including family and community characteristics, childcare and learning environments. LSAC data also includes linkages to administrative data such as Medicare usage, NAPLAN results, and AEDI scores.

Further information
Further information about the study, including data access, the data dictionary and links to publications and reports, is available at:
The Australian Longitudinal Study on Women’s Health—widely known as ‘Women’s Health Australia’—is a longitudinal population-based survey, which has examined the health of Australian women since it commenced in 1995.

Who is involved?
The project is designed to explore factors that influence health among women who are broadly representative of the entire Australian population. To date the study has involved over 40,000 women from three original age-cohorts (women born 1921–26; women born 1946–51; and women born 1973–78) as well as a new cohort of young women recruited in 2012 (women born 1989–94).

The study provides an evidence-base to the Australian Government Department of Health and Ageing for the development and evaluation of policy and practice in many areas of service delivery that affect women.

Content areas
The study goes beyond a narrow perspective that equates women’s health with reproductive and sexual health, and takes a comprehensive view of all aspects of health throughout women’s lives. The project assesses:

• physical and emotional health (including wellbeing, major diagnoses, symptoms)
• use of health services (GP, specialist and other visits, access, satisfaction)
• health behaviours and risk factors (diet, exercise, smoking, alcohol, other drugs)
• time use (including paid and unpaid work, family roles and leisure)
• sociodemographic factors (location, education, employment, family composition)
• life stages and key events (such as childbirth, divorce, widowhood).

Further information
Further information about the study, including data access, the data dictionary and links to publications and reports, is available at: http://www.alswh.org.au/
Understanding increases in obesity

Cross-sectional data from the ABS shows that over 50 per cent of women and nearly 70 per cent of men are overweight or obese—rates that are higher than only four years ago. The reasons for this increase, while complex and still relatively poorly understood, are an urgent focus of public health policy.

The increases among some women may be partly explained by recent research using data from the Australian Longitudinal Study of Women’s Health showing that employed women aged 45–50 tend to gain increasingly more weight over a two-year period the more hours they work. The findings suggest that women in this age group, who in the last decades have tended to work much longer hours in paid employment, have less time to maintain a healthy lifestyle than they did previously.

14 Australian Longitudinal Study on Male Health

The Australian Longitudinal Study on Male Health, known as the Ten to Men Study, is a longitudinal population-based survey which examines the health of Australian males. The study was funded under the 2010 National Male Health Policy.

Who is involved?
The Ten to Men Study is designed to examine the social, economic, environmental and behavioural factors that affect the length and quality of life of males aged 10 to 55 years. Three age cohorts will be examined (boys aged 10–15 years; youths aged 16–18 years; and adults aged 19–55 years). Males aged 10–18 years and those living in rural and regional areas will be oversampled.

The study is intended to provide an ongoing evidence-base for understanding the health and wellbeing of Australian males. Information from the study will be used to inform policy and practice in areas of service delivery that affect men, particularly groups of men at risk of poor health and at key life transition points.

The first wave of the study will commence in the field in early 2013.

Content areas
The study will involve the use of three age-appropriate survey instruments that focus on the following areas, as relevant to each cohort:

- health risk status (e.g. chronic and acute conditions, sexual and reproductive health, mental health, disability)
- risk and protective factors, including a range of social and environmental determinants (e.g. demographic characteristics, education, employment, socioeconomic status, social and gender roles)
- health risk behaviours (e.g. diet, exercise, smoking, drug and alcohol consumption, driving, safe sex)
- use of health goods and services (e.g. medical, dental, allied health and other health practitioners, use of non-prescription medications, Therapeutic Goods Administration listed drugs and therapeutic devices).

Further information
Further information about the study is available at: http://www.tentomen.org.au/
15 Longitudinal Surveys of Australian Youth (LSAY)

LSAY provides a rich source of information to help better understand young people and their transitions from school to post-school destinations, as well as exploring social outcomes, such as wellbeing.

Who is involved?
LSAY collects information about education and training, work, and social development from large, nationally representative samples of young people.

Survey participants enter the study when they turn 15 years, or as was the case in earlier studies, when they were in Year 9. Individuals are contacted once a year for 10 years. Studies began in 1995 (Y95 cohort), 1998 (Y98 cohort), 2003 (Y03 cohort), 2006 (Y06 cohort) and, more recently, in 2009 (Y09 cohort). Since 2003, the initial survey wave has been integrated with the Organisation for Economic Co-operation and Development (OECD) Programme for International Student Assessment (PISA). Over 10,000 students start out in each cohort.

LSAY is managed and funded by the Australian Government Department of Education, Employment and Workplace Relations (DEEWR), with support from state and territory governments. The LSAY data are collected by computer assisted telephone interviewing (CATI) undertaken by the Wallis Consulting Group.

Content areas
Information collected as part of LSAY covers a wide range of school and post-school topics, including: student achievement, student aspirations, school retention, social background, attitudes to school, work experiences and what students are doing when they leave school. This includes vocational and higher education, employment, job seeking activity, and satisfaction with various aspects of their lives.

Further information
Further information about the study is available at: http://www.lsay.edu.au/
**Which vocational education pathways lead to better employment outcomes**

Not surprisingly, students who complete year 12 schooling have better employment outcomes overall than those who don’t. However, for some students with strengths in non-academic skills, it may be better to complete another vocational education pathway after year 10 than staying on at school for another two years, for example by completing an apprenticeship or some other vocational certificate. Understanding whether all students will benefit from continuing to year 12 is therefore an important question for informing policy settings.

Research using data from the Longitudinal Surveys of Australian Youth (LSAY) found that for those students who run the risk of dropping out of school early, either completing an apprenticeship or finishing year 12 tends to lead to better employment outcomes. Without longitudinal data it wouldn’t be possible to track and compare these different pathways over time.

**Further Information:** Chris Ryan, Year 12 Completion and Youth Transitions, Longitudinal Surveys of Australian Youth: Research report no. 56, National Centre for Vocational Education Research, Adelaide, 2011.
**16 Longitudinal Study of Indigenous Children (LSIC)**

*Footprints in Time*, the Longitudinal Study of Indigenous Children, is an initiative of the Australian Government, conducted by FaHCSIA under the guidance of a steering committee, chaired by Professor Mick Dodson AM. The study aims to improve the understanding of, and policy response to, the diverse circumstances faced by Aboriginal and Torres Strait Islander children, their families and communities.

**Who is Involved?**

Since 2008, *Footprints in Time* data has been collected annually by Indigenous interviewers from a sample of over 1600 Aboriginal and Torres Strait Islander children and their families around Australia. The study includes 11 sites, including urban, regional and remote areas. The study involves two cohorts of Indigenous children with a target age from 6 months to 2 years (B cohort) and from 3 years, 6 months to 5 years (K cohort) at the time of the wave 1 collection.

**Content areas**

Data collected includes household information, child health, parent health, child and family functioning, socioeconomic and demographic information about the family, and assessment of the children’s development using a range of child outcome measures.

**Further information**

Further information about the study is available at: http://www.fahcsia.gov.au
17 Business Longitudinal Database (BLD)

The Business Longitudinal Database (BLD) is a valuable resource that will allow analysts to study the impacts of policies and industry trends on Australian businesses at the firm level. The database will enhance understanding of the activities or factors that are relevant to business performance; and the business characteristics that are associated with these activities or factors.

Who is involved?
The sample design uses panels that represent the Australian business population at the point in time that each panel is initiated into the BLD. Panel One is representative of the in-scope business population as at 30 June 2005. Panel Two is representative of the in-scope business population as at 30 June 2006. Each panel is directly surveyed once a year for a period of five years.

The database was developed and is managed by the ABS.

Content areas
Some of the key data items on the BLD confidentialised unit record file (CURF) include:

- general business characteristics—industry division, business size (based on employment), number of locations, whether a business is home-based and length of operation
- employment—total numbers, full or part time, casuals, pay setting arrangements
- basic financial indicators—value of sales and capital/non-capital purchases
- innovation indicators—type and status of innovative activity
- information technology indicators—use of internet, web presence, internet commerce
- market and competition—geographic markets, sources of income, customer relations, number of competitors and market share
- business financing—type, status and reasons for finance
- barriers to business performance.

Further information
Further information about the study is available at: http://www.abs.gov.au
18 Statistical Longitudinal Census Dataset (SLCD)

The 5 per cent Statistical Longitudinal Census Dataset (SLCD) brings together data from the 2006 Census with data from the 2011 Census and future censuses to build a picture of changes in society. The dataset offers the opportunity to investigate transitions of circumstances for both households and individuals.

Who is involved?
The longitudinal dataset links 5 per cent of records (approximately 1.1 million) from the 2006 and 2011 Census without using name and address but by using characteristics common to both datasets, such as age, sex, geographic region and country of birth.

The dataset is being developed and managed by the ABS.

Content areas
The SLCD covers all areas of Census data including:
- demographic information
- family structure
- household information
- labour force and education information
- caring responsibilities and disability information
- voluntary work.

The SLCD can be used to analyse transitions between Census years and provides supporting information for assessing the representativeness of other longitudinal survey data as well as future survey design.

Further information
The 5 per cent SLCD containing 2006 and 2011 Census data will be available for statistical analysis and research purposes from 2013. More information is available at: www.abs.gov.au
Longitudinal Study of Factors Affecting Housing Stability

Journeys Home: Longitudinal Study of Factors Affecting Housing Stability aims to improve the understanding of, and policy response to, the diverse social, economic and personal factors related to homelessness and the risk of becoming homeless.

Who is involved?
The research combines de-identified longitudinal information held by the Department of Human Services with a sample survey of approximately 1600 income support recipients across Australia. Individuals have been selected into the study if they are either homeless or at risk of homelessness. Journeys Home has collected four waves of data, each six months apart, from September 2011 to the first half of 2013.

The study was initiated, and is funded, by FaHCSIA. Responsibility for the design and management of the survey rests with the Melbourne Institute of Applied Economic and Social Research (the University of Melbourne) which sub-contracts Roy Morgan Research for the fieldwork. The Department of Education, Employment and Workplace Relations and the Department of Human Services are also partners in Journeys Home.

Content areas
Journeys Home collects important information on homelessness, covering areas such as:

- personal circumstances—participants’ physical and mental health; participation in the workplace, employment, education and training; and any significant life events
- family circumstances—participants’ family status and living arrangements, support networks and experiences of domestic and other violence
- housing circumstances—participants’ housing situation; the periods, nature of, and reasons for, homelessness; and trajectories of, and tipping points for, moving into and out of homelessness
- use of support services—types of assistance sought and used, including health care and support services.

Further Information
Confidentialised data will be made available for academics and other agencies for future research or policy development. The waves 1 and 2 research report have been published, as all future reports will be, at: http://homelessnessclearinghouse.govspace.gov.au

Further information on Journeys Home is available at: www.fahcsia.gov.au
Analysing the factors that support people at risk of homelessness

The Australian Government’s White Paper on Homelessness: The Road Home recognised the lack of information we have on the lives of the homeless and those at risk of homelessness over time.

- What are their characteristics?
- What factors are associated with instability/stability in housing tenancy or occupancy?
- How often and for how long do people experience homelessness?
- Why do some people become homeless while others in similar circumstances do not?
- What are important factors that may lead to the road out of homelessness?
- While Journeys Home is still in its relatively early stages, a picture is forming that will, when complete, inform homelessness policy and guide the strategic investment of government programs.
Appendix 1: Difference between repeated cross-sectional data and longitudinal data

Figure 1a: Repeated cross-sectional data.

Figure 1a shows a very simplified example of repeated cross-sectional data. This example considers changes in income over time. It involves collecting data from two individuals (sampled from a larger population) at three different points in time.

Importantly, for most repeated cross-sectional surveys, the individuals from whom the data are collected are unlikely to be the same on each occasion. Even if the same individual was sampled in each year, the data is not collected in a way that allows that person’s data to be linked and tracked over time.

However, the data does allow us to calculate the average income at each year, and then show how this increases over the period.
Figure 1b shows data that is numerically the same, but collected and analysed longitudinally. In this scenario data is gathered on each occasion from the same two individuals, Susan and David. Because each individual’s data records are linked over time, we can now see how each of their individual incomes changed. In this case, the data shows Susan’s income started lower than David’s but then overtook his by year 3, whereas David’s income increased slightly in year 2 and then decreased sharply by year 3. Only longitudinal data allows us to compare these individual trajectories over time. Of course, in the same way we showed with repeated cross sectional data, we can also see how the overall average of Susan’s and David’s income increases over the period too.
## Appendix 2: Comparison of different types of data

### Table 1a: Applicability of type of data to types of analysis

<table>
<thead>
<tr>
<th></th>
<th>Longitudinal data</th>
<th>Cross-sectional data</th>
<th>Randomised control trials</th>
<th>Qualitative research, e.g. focus groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demonstrate causal processes</td>
<td>✓✓</td>
<td>✓</td>
<td>✓✓✓</td>
<td>-</td>
</tr>
<tr>
<td>Evaluate diverse outcomes over time</td>
<td>✓✓✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Quantify persistence and transition</td>
<td>✓✓✓</td>
<td>✓</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Quantify prevalence</td>
<td>✓✓</td>
<td>✓✓✓</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Qualitatively describe individual stories</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>✓✓✓</td>
</tr>
</tbody>
</table>

### Key

- **✓✓✓** best available standard of evidence, where practically and ethically feasible
- **✓✓** good scope for conducting this analysis – but some limitations
- **✓** limited scope to conduct this analysis
- **-** rarely or never used or designed for this purpose
Glossary

Cohort: a group of individuals who share a fixed characteristic or have shared an event during a particular time span, for example all children born in a particular year.

Cross-sectional data: collected by observing many subjects (such as individuals or households) at the same point in time (or over the short period of data collection)—analysis of cross-sectional data usually consists of comparing the differences between subjects.

Longitudinal data: collected from the same individuals, or households, at several points in time—the data is designed so that differences within and between individuals can be tracked over time.

Repeated cross-sectional data: the same data items are measured using separate samples of the population taken at two or more points in time—the data does not normally measure the same individuals at each time point. For this reason, this type of data can measure aggregate changes in a population but cannot measure changes within individuals over time (as with longitudinal data, described above).

Time series data: a type of repeated cross-sectional data. A time series is a collection of observations of well-defined data items obtained through repeated measurements over time. For example, measuring the level of unemployment each month of the year would comprise a time series. This is because employment and unemployment are well defined, and consistently measured at equally spaced intervals. Data collected irregularly or only once are not time series. As with other repeated cross-sectional data designs, time series are unlikely to measure the same individuals at each time point and so only aggregate change, not change within individuals, can be measured.

Wave: a period of data collection, for example, in LSAC, wave 1 occurred in 2004 when families were interviewed for the first time; Wave 2 occurred in 2006 when the same families were interviewed for the second time.