



**YOUTH IN FOCUS PROJECT DISCUSSION PAPER SERIES
No. 8, December 2009**

User's Guide to the Youth in Focus Data
Version 2.0

**Robert Breunig
Deborah Cobb-Clark
Tue Gørgens
Chris Ryan
Anastasia Sartbayeva**

youthinfocus.anu.edu.au



Jointly funded by the Australian Government Department of Education, Employment and Workplace Relations, the Australian Government Department of Families, Housing, Community Services and Indigenous Affairs and the Australian Research Council, and carried out by The Australian National University

Youth in Focus Project Discussion Paper Series
ISSN 1835-4025

Youth in Focus
Research School of Social Sciences
ANU College of Arts and Social Sciences
Coombs Building 9
The Australian National University
Canberra ACT 0200

Table of Contents

Overview	4
1. Introduction	5
1.1. Project organisation and partners	5
1.2. Project motivation	5
1.3. Main research questions.....	7
1.4. Data sources	7
1.5. Data management, access and security.....	8
2. Reference populations and project terminology	9
2.1. Birth cohort and focal youth	9
2.2. Parents.....	9
2.3. Observational unit.....	10
2.4. Data collection unit	10
3. Administrative data and sampling frame	11
3.1. The administrative database (TDS2)	11
3.2. TDS2 and the Australian population.....	12
3.3. Currency of contact information.....	13
4. Survey design	13
4.1. Data collection mode	13
4.2. Interview frequency	13
4.3. Payment to respondents.....	13
4.4. Sample stratification	14
4.5. Sample size in wave 1	14
4.6. Sample size in the following waves.....	16
5. Survey instruments	16
5.1. Wave 1 instruments	16
5.2. Wave 2 instruments	17
6. Fieldwork	17
6.1. Contracting of market research firm	17
6.2. Privacy and confidentiality issues	18
6.3. Interviewer monitoring and complaint handling protocol.....	18
6.4. Refusal conversion and follow-up.....	18
6.5. Wave 1 fieldwork	18
6.5.1. <i>Pre-test</i>	18
6.5.2. <i>Interviewer training</i>	19
6.5.3. <i>Wave 1 pilot</i>	19
6.5.4. <i>Wave 1 pilot opt-out process</i>	19

6.5.5. Wave 1 pilot outcomes	20
6.5.6. Wave 1 sample.....	20
6.5.7. Wave 1 opt-out process	20
6.5.8. Wave 1 data gathering procedure	22
6.5.10. Interview length and outcomes of self-completion questionnaire	24
6.5.11. Timeline.....	24
6.5.12. Payment to respondents	24
6.5.13. Foreign language interviews.....	25
6.5.14. Wave 1 contact outcomes	25
6.5.15. Data linkage	27
6.5.16. Tracking and tracing.....	28
6.6. Wave 2 fieldwork	28
6.6.1. Interviewer training.....	28
6.6.2. Wave 2 pilots.....	28
6.6.3. Wave 2 pilot outcomes	29
6.6.4. Wave 2 sample.....	29
6.6.5. Wave 2 data gathering procedure	30
6.6.6. Interview length and outcomes of self-completion questionnaire	30
6.6.7. Timeline.....	31
6.6.8. Payment to respondents	31
6.6.9. Foreign language interviews.....	31
6.6.10. Wave 2 contact outcomes	31
6.7. Future waves.....	33
7. Survey dataset.....	34
7.1. Variable naming conventions.....	34
7.2. Labeling	35
7.3. Survey datafiles characteristics	36
Appendix A. Administrative dataset.....	37
Appendix B. YIF questionnaires summary	41
B1. Parent questionnaire.....	41
B2. Youth wave 1 questionnaire.....	42
B3. Youth wave 1 self-completion questionnaire	44
B4. Continuing respondent (youth) wave 2 questionnaire	45
B5. New entrant (youth) wave 2 questionnaire	46
B6. Youth wave 2 self-completion questionnaire	48
Appendix C. Additional information on wave 1 pilot.....	49
C1. Items tested in pilot.....	49
C2. Pilot response rates.....	50
Bibliography.....	52

List of tables

Table 1.	Number of young people in TDS2 and the Australian population.....	12
Table 2.	Proportion of people across income support stratification categories.....	15
Table 3a.	Wave 1 sample and mailout structure for focal youth.....	22
Table 3b.	Wave 1 sample and mailout structure for parents.....	22
Table 4a.	Wave 1 focal youth outcomes.....	26
Table 4b.	Wave 1 parent outcomes.....	27
Table 5.	Respondents' consent to data linkage, per cent of total.....	27
Table 6a.	Wave 2 continuing respondents outcomes.....	32
Table 6b.	Wave 2 new entrants outcomes.....	33
Table 7.	Survey dataset introductory variable description.....	34
Table 8.	Survey datafiles characteristics.....	36
Table A1.	Total numbers of focal youth in the administrative dataset (main birth cohort) by gender, month of birth and stratification category.....	37
Table A2.	Total numbers of primary parents in the administrative dataset by gender and stratification category.....	39
Table C2.1	Wave 1 pilot response rates.....	49
Table C2.2	Wave 2 pilots response rates.....	50

Overview

This paper provides important background information regarding the Youth in Focus project, and the associated survey and administrative data. We outline the motivation behind the project, the primary research questions that the project is intended to address, the funding sources, and the composition of the project research team.

Furthermore, we provide detailed information on the methodology of the project. This includes the data sources, the survey sampling frame, the sampling methodology, a brief description of the survey instruments, and the timeline of fieldwork. We describe the pilot study that was conducted, the use of respondent payments in the project, the survey outcomes from the pilot and the first two waves of data collection.

This version of the User's Guide provides an update on wave 1 procedures and datasets, and describes wave 2 survey methodology, outcomes, and datasets.

1. Introduction

1.1. Project organisation and partners

Youth in Focus (YIF) is a joint research project between the Australian Government Department of Education, Employment and Workplace Relations (DEEWR)¹ and a team of academic researchers. The current research team includes Dr Robert Breunig, Professor Deborah Cobb-Clark, and Drs Tue Gørgens and Chris Ryan of the Australian National University (ANU), Professor Jeff Borland of the University of Melbourne, and Professors Barbara Wolfe and Robert Haveman of the University of Wisconsin-Madison.²

The YIF project is supported by a five-year (2003-2008) Linkage-Project grant from the Australian Research Council (ARC).³ In addition to its initial commitment of resources for the project, the Department of Families, Housing, Community Services and Indigenous Affairs (FaHCSIA) and DEEWR have provided support for background research relevant to the project. Centrelink has given in-kind support by providing staff time and expertise for work related to building the administrative data set. The ANU is the lead institution and has primary responsibility for coordinating the work of the research team, DEEWR, and Roy Morgan Research Pty Ltd (RMR) – the market research firm conducting the survey – and for reporting to the ARC.

1.2. Project motivation

The overarching goal of the Youth in Focus project is to understand the ways in which economic and social disadvantage might be transferred from one generation to the next. The project uses an innovative combination of survey and administrative data to explore some of the consequences for young Australians of growing up in disadvantaged families. Our focus is on outcomes in the early adult years when young people are moving into higher education, entering the labour market, starting families, and generally establishing themselves as independent adults. Developing a fuller understanding of reasons why disadvantaged youth succeed – or fail to

¹ From inception to May 2008, the Commonwealth Government partner in YIF was the Department of Families, Housing, Community Services and Indigenous Affairs (FaHCSIA). Until 2006, it was known as the Department of Families and Community Services (FaCS). Two Australian Government reorganisations resulted in FaCS becoming first the Department of Families, Community Services and Indigenous Affairs (FaCSIA) on 27 January 2006 and then the Department of Families, Housing, Community Services and Indigenous Affairs (FaHCSIA) on 3 December 2007. Following administrative changes implemented by the Australian Government in late 2007, responsibility for YIF passed to DEEWR, effective May 2008.

² From inception to 2006, the research team also included Ms Jocelyn Pech of Department of Employment and Workplace Relations (previously with Department of Family and Community Services).

³ARC Linkage Project LP0347164 entitled 'The Intergenerational Transmission of Dependence on Income Support: Patterns, Causation and Implications for Australian Social Policy Research'.

succeed – is a necessary first step in formulating sensible policies targeted towards breaking any cycle of dependence and promoting the social and economic independence of Australian youth.

In particular, it is well-established that individuals who grow up in families that are dependent upon income support are themselves much more likely to be dependent upon income support as adults.⁴ At the same time, the majority of children who grow up in income-support dependent families are not reliant upon income support as adults. An important objective of the project is identifying the factors which contribute to or prevent the need for social assistance amongst young adults.

Although the association between growing up in an income-support dependent family and reliance upon social assistance as an adult are well-established, there has been only limited research on assessing which factors underlie this relationship. A lack of educational qualifications, early childbearing, poor health and disability all contribute to reducing a young person's labour market opportunities and increasing the chances of needing social assistance.⁵ Determining the causal effect of these risk factors is difficult, since while all can cause dependence on income support, each may also result from a history of dependence on income support.

Previous studies have faced two major empirical challenges. Most studies have been based upon a population of disadvantaged individuals receiving support from a particular government program. These studies have not had a comparison group of program non-participants to provide a compelling counterfactual to answer questions such as 'does dropping out of school early make youth in welfare-dependent families more likely to access social assistance as adults than youth from non-welfare-dependent families who drop out of school early?' The second difficulty that previous studies have faced is the lack of complete administrative data about program participation. These studies have had to rely on individuals' own reports of income-support receipt and have usually compared two particular points in time – one during an individual's youth and one in adulthood. In addition to raising measurement issues that restrict the possible observable associations and complicate interpretation,⁶ such studies are unable to say anything about the importance of timing or intensity of exposure to social assistance.

This project is designed to address and overcome the weaknesses of previous studies. By using a combination of administrative data going back to 1991 and survey data gathered from both parents and children, issues of timing, intensity, and incidence of disadvantage can be studied while controlling for a range of background and demographic factors. The nature of the Australian payments system, managed by

⁴ See, for example, Pech and McCoull (1998) and McCoull and Pech (2000) for Australian evidence and Solon et. al. (1991), Gottschalk (1996), Haveman and Wolfe (1995), Israel and Seeborg (1998), and Haveman et. al. (2001) for international evidence.

⁵ See Haveman et. al. (2001).

⁶ See An et. al. (1996).

Centrelink, provides a population sampling frame of young Australians who grew up in a range of family circumstances. In particular, the same administrative database used to manage income-support payments is also used to manage childcare subsidies (which are not means tested) and tax rebates for dependent children (which are means tested and not paid to the top 15 per cent (approximately) of the income distribution). Thus the data source from which the survey sample is drawn provides consistent administrative data not only for disadvantaged families, but also for a large comparison group of middle and upper-middle income families.

1.3. Main research questions

The data generated by the project will be the foundation for research on a wide variety of issues including youth health, education and employment, as well as the influence of intergenerational factors on young people's socio-economic outcomes. The overarching goal of the YIF project is to understand the consequences of growing up in disadvantage on economic, social, and demographic outcomes in early adulthood.

To this end, the research project has two main objectives. The first objective is to describe the overall correlation between parents' and young adult children's outcomes. The second objective is to investigate the causal mechanisms behind these relationships, paying particular attention to the intergenerational correlation in income-support receipt. The literature points to several mechanisms through which disadvantage might be passed from one generation to the next however, little is known about their relative importance. Identifying which mechanisms are most important in accounting for the intergenerational correlation in disadvantage is a necessary first step in formulating evidence-based policies targeted towards breaking any cycle of welfare dependence.

Data collection is designed to focus on four possible transmission mechanisms: low educational attainment, early fertility, poor health and/or disability, and attitudinal factors. All of these may depend upon the socio-economic status and income-support history of a young person's family and may affect a youth's future receipt of income support. Information gathered from matched pairs of young adults and parents, lengthy administrative data, and a two-wave panel survey provide the mechanisms by which these research questions may be addressed.

1.4. Data sources

There are two data sources for the project. The first is administrative data collected by the Australian Government (Centrelink) from all recipients of (and applicants for) any type of government payment. The second data source is a survey of a sample of individuals selected from the administrative data specifically for the YIF project.

The administrative data contain information about most payments from the Australian Government to Australian families. It is believed to be nearly

representative of the entire Australian population, with the exception of families with very high income. According to FaHCSIA 2004-2005 estimates, approximately 85 per cent of families with children receive Family Tax Benefit and therefore appear in the administrative data.⁷ A smaller number of families receive additional support primarily due to low income. (Disability status could be a reason for income-support receipt unrelated to low income.) This support may take the form of unemployment benefits, payments to single parents with low income, payments to couples with young children and low income, and a variety of less common payments.⁸ Since the administrative data include families who received substantial income support as well as families who received only the tax benefit, the administrative data provide an excellent basis for studying the correlation between parents' and children's receipt of income support from a combined sample of recipients and non-recipients.

The survey is designed to collect a variety of information from a random sample of families who have appeared in the administrative data at least once since 1991. The survey design is based around a birth cohort of youth who turned 18 just before wave 1 interviewing (see Section 2.1 below). For each youth, we identified all individuals who received any type of payment on behalf of that youth in the history of the administrative data. From these individuals, who we henceforth refer to as 'parents', we selected the person who had the longest duration of care (as measured by receiving a payment on behalf of the youth or claiming the youth as a dependent) over the history of the administrative data set. We further implemented a set of rules designed to identify the natural mother from the administrative data. Although the administrative data do not contain information on the actual family relationships, the chosen method proved extremely successful in identifying the natural parents. Among wave 1 respondents, a natural mother was selected in 96.5 per cent and a natural parent in 98.6 per cent of cases.

The parents were interviewed once and at the time of writing the children were interviewed twice with a gap of approximately two years between the waves. Respondents were asked to provide information on topics such as employment, education, physical and mental health, attitudes and values, family relationships and other psycho-social factors, the children's experiences while growing up, neighbourhood and school quality.

With the respondents' consent, the survey data were linked to the administrative data, and the linked data form the basis for studying the causal mechanisms.

1.5. Data management, access and security

Due to the sensitivity of the confidential data contained in the administrative dataset,

⁷ Tax benefits for dependent children can take the form of a tax deduction or of a bi-weekly payment direct to families. The vast majority of families in Australia who are eligible for Family Tax Benefit opt for the bi-weekly payment.

⁸ See Centrelink (2007).

its access at the ANU is restricted to a secure data room. The secure data room and procedures for using the data in it incorporate a range of measures designed to protect the data. This room and the procedures are regularly audited and inspected. It is anticipated that the survey data will be made accessible in the future to other interested researchers, together with a summary of information from the administrative dataset.

The main management of the project has been done through the project steering committee. This committee met several times per year and was composed of members of the project research team and representatives from DEEWR and FaHCSIA. This committee was responsible for dissemination of the data, designing conditions of use for the data, and approving applications to use the data.

2. Reference populations and project terminology

2.1. Birth cohort and focal youth

The primary reference population for the main wave of the YIF survey is all people born between 1 October 1987 and 31 March 1988 who appear in the Centrelink administrative data at any point between 1991 and July 2006. We will refer to this group as the 'main birth cohort'. In addition, there is a smaller group of young people born in July 1987 and January 1988 only, which is used in the pilot stages of the survey.⁹ We will refer to these young adults as the 'pilot birth cohort'. For both cohorts, the selected young person in each family will be referred to as the 'focal youth'.

There are two ways that a young person can appear in the administrative data. One is if the young person has received a payment, such as Youth Allowance, directly from the Australian government in his/her own name. The other is that the young person has been listed as the dependent child of another individual for the purpose of obtaining eligibility and amount of payment such as Family Tax Benefit or Parenting Payment.

For the survey purposes, certain focal youth participants are considered out-of-scope (for details, see Section 6.5.8). These included, for example, those who were in the armed forces, overseas, or institutionalised with no prospect of return until the end of the survey.

2.2. Parents

The secondary reference population includes all adults who, at any point since 1991, have received government payments for the focal youth, or who had the focal youth registered as their dependent for payment purposes. The vast majority of cases will

⁹ The two months were originally selected for use in the pilot to test for any differences in the contactability of focal youth. The pilot has shown no significant difference in response rates.

be where the adult received Family Tax Benefit (previously Family Allowance), which is the annual tax benefit to help families with the costs of raising children. The adult in question may be a birth parent, an adopted parent, a foster parent, an older sibling, a relative, or a non-relative who cared or is caring for the focal youth. Although the family relationship between the focal youth and these adults is not identifiable from the administrative data, we will refer to these people as 'parents'.

Due to the construction of the parent subset of the administrative data, a focal youth may have several parents or guardians. For the purposes of the survey, only one parent was selected, the criterion being the longest duration of care (payment receipt) for the focal youth. By looking at the gender of parents and the timing of receipt from the administrative data, we attempted to select the natural mother whenever possible. As mentioned previously, this methodology resulted in the natural mother being selected in 96.5 per cent, and a natural parent being selected in 98.6 per cent of cases.

As with focal youth, certain parents are considered out-of-scope (for details, see Section 6.5.8).

2.3. Observational unit

Most of the focal youth have a parent in the administrative dataset. However, in some cases there is no parent. This may happen if the parent(s) are deceased or if the focal youth is in the administrative data as a recipient of Youth Allowance and his/her parent(s) have never received any type of Centrelink payment.

In cases where both the focal youth and a parent are available in the administrative data, the matched focal youth-parent twosome will be referred to as the 'observational unit'. In cases where there is no parent record, the observational unit is the focal youth. Each observational unit, therefore, has either one or two members, depending upon whether there is a record for a parent. Note that the observational unit is not necessarily a complete family nor a household. The proportion of one person (focal youth only) observational units in the administrative data is 1.7 per cent, with the two person (focal youth and parent) observational units constituting the overwhelming majority at 98.3 per cent.

2.4. Data collection unit

The data collection unit is the individual. Different instruments are used for data collection from parents and youth, as detailed below.

3. Administrative data and sampling frame

3.1. The administrative database (TDS2)

The frame for this study is an administrative database (referred to as TDS2¹⁰) built from Centrelink payment records. Centrelink represents a one-stop shopfront for receipt of almost all government benefits including parenting and family payments, unemployment benefits, old-age and disability pensions, and provides an excellent sampling frame because most transfer payments are administered through this one agency at the national level.

All focal youth who appeared in the Centrelink payment records at any time since 1991 are included in this database. A focal youth may appear in the administrative data for two reasons: (i) they have claimed in their own right Centrelink payments, such as Youth Allowance or disability payments; or (ii) at some point since 1991, they were listed as a dependent of an adult who has received a government payment. Any adult who claimed payments at any time for the focal youth is also included and is referred to as a parent. As mentioned earlier, the relationship between parent and the focal youth is an administrative one and thus the adult may not be a natural parent.

In addition, the dataset includes 'siblings', identified as other children for whom the parents of the focal youth also claimed payments. Again, these siblings may have no blood relation to the focal youth. Sibling records are linked to the parent records and through the parent to the focal youth records.

For all people in the administrative dataset who received payments in their own right, the following information is provided:¹¹

- Personal details: date of birth, sex, country of birth;
- Housing details: home postcode history, accommodation history, rent type;
- Youth homeless or independent¹² history;
- Full time student history;
- Family income-support history;
- Marital status history;
- Individual income-support history (payment types and dates of receipt).

¹⁰ The first Transgenerational Dataset (TDS1) is a database that was created by FaHCSIA in 1996 and informed the construction of TDS2.

¹¹ Since the responsibility of providing the updated information lies with the users of Centrelink, some personal information may not be complete or regularly updated. Moreover, there is no requirement for individuals to update their information once they stop receiving benefits. 'TDS User guide' written by Centrelink (Wellman, 2005) contains more details on information available in TDS2.

¹² In certain circumstances, young adults (over 15 years of age) may receive an 'Independent Status' and be considered independent from their parents for the purposes of Centrelink payments receipt.

The entire observational unit was eliminated from the frame if the administrative data identified any member of the observational unit as falling into one of the following categories:

- Their administrative record is marked 'Deny Access', which is a facility available to people seeking additional protection of personal information held by Centrelink, for instance, where there has been or is a potential of domestic violence;
- They have indicated refusal to participate in any surveys;
- Focal youth is deceased.

3.2. TDS2 and the Australian population

The numbers of focal youth present in the administrative data can be compared against other available statistics on young people born between 1 October 1987 and 31 March 1988. Table 1 shows the number of youth records in the administrative database (TDS2), as well as the number of births registered during the relevant months¹³ and the Census 2006 results¹⁴. The 2006 Census data refer to the proportion of 263,222 18-year olds (excluding overseas visitors) located in Australia on the Census night of 8 August 2006. The applied proportion of 49.59 per cent is the number of births during October 1987 to March 1988 relative to the number of births which occurred during August 1987 to July 1988, as reported in ABS (2002, 2003).

Table 1. Number of young people in TDS2 and the Australian population

	TDS2	Registered births	Census 2006 ^a
Born October-December 1987	62,801	59,863	–
Born January-March 1988	64,999	61,844	–
Total	127,800	121,707	130,532

Notes:

a The Census data show the proportion of all 18-year olds (excluding overseas visitors) located in Australia on the Census night of 8 August 2006. Since the Census data do not provide information on the month of birth, the applied proportion of 49.59 per cent is derived based on the proportion of births in the birth cohort months (October 1987 to March 1988) relative to births occurred during August 1987 to July 1988.

Sources:

ABS (2002, 2003), Census 2006 (available from ABS website at <http://www.abs.gov.au>).

¹³ See ABS (2002 and 2003).

¹⁴ <http://www.abs.gov.au/websitedbs/d3310114.nsf/Home/census>

Table 1 demonstrates that the administrative dataset contains records on nearly all young adults belonging to the birth cohort. The number of registered births does not provide up-to-date information, since it does not take into account migration and deaths.

3.3. Currency of contact information

Centrelink customers are required to provide their contact details when they first apply for payments and are required to inform Centrelink of any changes in their circumstances. However, since payments do not necessarily depend on the availability of current contact details (payments usually being made directly into specified bank accounts), customers may neglect to keep their records current. Moreover, individuals are not required to update personal records if they no longer receive any payment.

In view of this, individuals with short exposure to the income-support system which took place a long time ago may be more difficult to locate than those with more recent exposure. To compensate for this, such individuals were over-sampled (see Section 4). In addition, multiple approaches were employed to find missing contact details (detailed in Section 6).

4. Survey design

4.1. Data collection mode

The primary data collection mode is computer-assisted telephone interviewing (CATI). In addition to the questionnaire administered by telephone, the focal youth respondents are asked to complete a self-completion questionnaire. Youth may choose to complete this on-line at a secure site or to complete a paper questionnaire which is returned by mail.

4.2. Interview frequency

Due to the longitudinal nature of the YIF project, the focal youth were interviewed at two waves, conducted approximately two years apart. The database of contact details of the survey participants is maintained by the market research firm to ensure the possibility of gathering additional data if funding for future waves becomes available.

4.3. Payment to respondents

Based upon a test of payment to respondents in the pilot (see Section 6), it was decided to pay a small monetary compensation to all participants. Each participant who completed the interview received a cheque for \$15. Focal youth in waves 1 and 2 were required to also complete and return the self-completion questionnaire in order

to receive payment¹⁵. It is important to note that the payment was not based upon responses to any particular question. In fact, a participant could agree to participate, but subsequently refuse to answer every question and the compensation would still be paid.

4.4. Sample stratification

The sample of focal youth participants and their parents/guardians was selected as a probability sample, stratified on the basis of the income-support history of the parent as described in Table 2.

The ‘no parental income-support history’ category combines those observational units where there is no parent record along with those (the vast majority) where the parent has received only Family Tax Benefit since 1993. Each focal youth will thus belong to one of the six stratification categories shown in Table 2. Appendix A provides more detailed information on the total numbers of youth and primary parents in the administrative dataset.

Table 2 shows that individuals belonging to strata B through F have been over-sampled in comparison to the proportions observed in the administrative data. Individuals with heavy exposure to the income-support system were over-sampled since they have been found in other surveys to have lower response rates relative to the average for the population. Another over-sampled group is people who have had only a short and not recent exposure to the income-support system, because of the anticipated difficulty in locating them relative to those with more recent exposure (see Section 3.3).

As the survey is conducted over the phone, it was decided not to stratify by geographic location.

4.5. Sample size in wave 1

We selected a sample of 13,568 young adults for wave 1 with the objective of having 4,000 completed interviews from youth, 3,797 completed interviews from parents, and 2,240 completed interviews from observational units.

The achieved final survey sample in wave 1 was 3,964 parent interviews and 4,079 focal youth interviews (refer to Section 6.5.14 for more detail).

¹⁵ For the main fieldwork of new entrants in wave 2, the payment to respondents was raised to \$25 if they completed and returned the self-completion questionnaire. Those who participated in the telephone interview only received a payment of \$15.

Table 2. Proportion of people across income-support^a stratification categories

Stratification category	Stratum code	Proportion in admin. dataset	Expected proportion in wave 1 survey dataset	Achieved proportion of youth in wave 1 survey data	Achieved proportion of parents in wave 1 survey data	Achieved proportion of matched obs units in wave 1 survey data
No parental income-support history	A	40.9%	25.0%	25.2%	26.4%	29.0%
Heavy exposure to income-support programs (more than six total years on income support)	B	27.5%	34.9%	36.1%	35.7%	32.3%
First exposure to income-support system after 1998	C	8.5%	10.7%	12.9%	12.4%	12.4%
First exposure to income-support system between 1994 and 1998 and less than three total years on income support	D	8.5%	10.8%	10.3%	10.6%	11.2%
First exposure prior to 1994 and less than six total years on income support	E	9.5%	12.1%	9.9%	9.7%	9.6%
First exposure to income-support system between 1994 and 1998 and more than three but less than six total years on income support	F	5.1%	6.5%	5.7%	5.4%	5.6%

Notes:

^a In this table, the term 'income support' refers to any kind of Centrelink benefits the parents may have received excluding family payments, Carer Allowance and maternity payments.

4.6. Sample size in the following waves

From the 4,007 focal youth who responded in wave 1 and were deemed in-scope for wave 2¹⁶, 2,362 interviews were collected. Thus the rate of attrition between wave 1 and 2 was slightly over 41 per cent. Such rate of attrition, although higher than originally expected, is comparable to that observed for this age group in other surveys (for instance, HILDA¹⁷).

In order to boost the survey sample, it was decided to attempt interviewing those wave 1 participants who were uncontactable at wave 1 or gave a 'soft' refusal to complete the wave 1 interview. This group of the survey participants will be referred to as 'new entrants' while those wave 2 respondents who have completed the wave 1 interview are referred to as 'continuing respondents'.

5. Survey instruments

5.1. Wave 1 instruments

Development of the survey instruments for wave 1 was an iterative process involving discussion between the project research team, FaHCSIA staff, the TDS2 development team, external technical advisors and other interested agencies. The teams working on the questionnaires drew on existing international and domestic surveys as well as expert advice from a range of individuals, including Ms Nicole Watson and Professor Mark Wooden at the Melbourne Institute of Applied Economic and Social Research who have responsibilities for HILDA as well as Professor Bryan Rodgers and Dr Helen Berry of the National Centre for Epidemiology and Population Health at the ANU. Professor Peter Lynn of the Institute for Social and Economic Research at the University of Essex, UK, also provided helpful advice.

The survey instruments were put into operation with the help of Roy Morgan Research Pty Ltd, the market research firm contracted to carry out the fieldwork. This involved pre-testing the questionnaires with several paid participants in October 2005, intensive practice runs of the survey instruments by the research team via a telnet access system to the CATI questionnaire, and a full-scale pilot of wave 1 of the survey in March-April 2006.

Three instruments were used in the YIF survey:

¹⁶ Wave 1 focal youth respondents were deemed out-of-scope for wave 2 if they were institutionalised, unwell, overseas or in the armed forces at the time of wave 2 (revealed either during the keep-in-touch program or at wave 2 contact).

¹⁷ Living in Australia: The Household, Income and Labour Dynamics in Australia Survey; more information is available at <http://www.melbourneinstitute.com/hilda/>

- *Parent questionnaire* which collected information on the parent's family, education, employment and income, relationship history, general values and health, as well as background information about the focal youth at early stages of life;
- *Youth questionnaire* which collected information on the focal youth's household and family, education, employment and job search, housing arrangements, income and health, as well as personality traits and life satisfaction;
- *Self-completion questionnaire for youth* which collected information on attitudes, substance abuse, relationships with parents and important life events.

The summary of topics covered by the survey instruments is provided in Appendix B. The full versions of the survey questionnaires are available from the research team upon request.

Where possible, wording from other Australian surveys was used in order to facilitate comparison of the data collected. Standard Australian Bureau of Statistics questions for items such as Indigenous status were employed. Questions relating to income and labour status were made similar to those in HILDA for comparability.

5.2. Wave 2 instruments

For the second wave of the survey, two sets of CATI questionnaires were developed: one for continuing respondents and one for new entrants. In the questionnaire for continuing respondents, we have deleted many of the items which relate to time-invariant information, such as migrant status of the focal youth and their childhood living arrangements, and added questions on youth's expectations of their situation in ten years' time and on their opinions about a range of current policy issues. The questionnaire for the new entrants was an amalgamation of wave 1 and wave 2 youth questionnaires, retaining questions on the youth's family background and including questions introduced in wave 2.

The same wave 2 self-completion questionnaire was used for both the continuing respondents and the new entrants. This questionnaire asked questions on attitudes, substance abuse and life events similar to those asked in wave 1, and a new set of questions on attitudes towards risk.

6. Fieldwork

6.1. Contracting of market research firm

A competitive tender process was undertaken to find a suitable organisation to undertake the fieldwork. It was anticipated that the chosen firm would assist in finalising survey instruments, designing and implementing a pre-test, and designing and implementing all waves of data gathering.

An information session was conducted at the Australian National University on 2

May 2005 for interested companies prior to the tender deadline. This was attended by representatives of six different companies. A phone information session was conducted for two companies which could not attend the information session.

Five proposals were received. From these, Roy Morgan Research (RMR) was selected as having the proposal which met the minimum requirements of the tender and which additionally offered the best value for money. The contract for provision of field services between RMR and the ANU was signed on 5 December 2005.

6.2. Privacy and confidentiality issues

The procedures for handling confidential data were set out in the YIF Project Ethics Protocol which was approved by the ANU Human Research Ethics Committee on 11 February 2005.

Due to the sensitive nature of the administrative data, all project personnel, including all interviewers and supervisors at RMR, have had to sign Deeds of Confidentiality relating to the treatment of personal contact and other information.

6.3. Interviewer monitoring and complaint handling protocol

The complaint handling protocol for the YIF survey was approved by the ANU Human Research Ethics Committee. In the case of complaints or queries not satisfactorily addressed by the interviewers, survey participants were referred to the interviewers' supervisors, followed, if necessary, by the project manager at the market research firm or the designated FaHCSIA/DEEWR representative.

6.4. Refusal conversion and follow-up

The research team, government organisations involved and the market research firm engaged to undertake the survey ensured that the participants were given the fullest information about the project in order to convert as many refusals as possible. Particular care was taken to emphasise the academic research nature of the survey and the government support of the YIF project to encourage participation.

Due to the longitudinal nature of the survey, the focal youth respondents were contacted several times between the waves in order to minimise attrition. The contacts were evenly spaced between waves and included a 'thank you' card, newsletter with a summary of the survey results, and a phone call closer to the next wave (refer to Section 6.5.16 for more details).

6.5. Wave 1 fieldwork

6.5.1. Pre-test

A small pre-test was conducted in October 2005 by RMR and the project research team with several paid participants to provide preliminary feedback on all survey instruments. As the main objective of the pre-test was to judge participants' reaction

to the questions asked, their answers were not recorded.

6.5.2. Interviewer training

In order to familiarise interviewers with the survey instruments and provide them with the background information on the YIF project, interviewer information sessions, led by the RMR project manager for YIF, were run immediately prior to the pilot and the main wave of interviewing. The pre-pilot information session was held on 27 March 2005 and was attended, on the ANU side, by Drs Robert Breunig and Tue Gørgens. Another session was held on 15 August 2006, immediately before the main fieldwork for wave 1, and was attended by Dr Robert Breunig, Professor Cobb-Clark and Dr Anastasia Sartbayeva as well as the representatives of the FaHCSIA Youth Bureau and the Research and Analysis Branch.

6.5.3. Wave 1 pilot

A pilot test was carried out between 27 March and 18 April 2006. The sample for the pilot consisted of observational units as described in Section 2.3, with the focal youth randomly selected from the pilot birth cohort.

To test for differences in response rates across stratification categories, the pilot sample contained observational units from the six strata in approximately equal proportions. In addition, to investigate the possibility of introducing respondent payments to encourage participation, each stratum was randomly divided into two halves, one of which was offered a \$15 compensation payment while the other was not offered any monetary compensation (see also Section 6.5.12).

6.5.4. Wave 1 pilot opt-out process

After drawing the initial raw sample of about 1,400 observational units, Centrelink provided FaHCSIA with contact information for the selected sample, which included name of focal youth, name of parent, and for each of these individuals a phone number (where available) and address.

Before passing the sample to RMR, FaHCSIA sent a primary approach letter and information brochure¹⁸ to each member of the observational unit. These materials contained sufficient information to enable people selected in the sample to reach an informed decision on whether or not they wished to participate in the survey.

Those who did not wish to participate were requested to call a 1800 number managed by FaHCSIA to opt out of the survey. Survey respondents could also use this number for any queries about the survey. After the end of the opt-out period, the 1800 number was transferred to RMR who then took the responsibility for respondent liaison.

¹⁸ The letter and information brochure can be viewed at <http://youthinfocus.anu.edu.au/survey.htm>

6.5.5. Wave 1 pilot outcomes

The information gathered in the pilot was used to judge the respondents' reaction to the questions, the effect of payment to respondents on response rates, and the accuracy of the CATI programming particularly as it related to questionnaire routing procedures. The full list of items tested in the pilot is available in Appendix C1.

After the interviews for the pilot were completed, RMR submitted a technical report on the outcome of the pilot study to the ANU. The report critically considered strategies and methodologies tested in the pilot study and made recommendations regarding proposed changes.

Analysis of the pilot data and the technical report resulted in the shortening of the parent questionnaire, correction of several glitches in the CATI programming, and a positive decision to use monetary compensation to improve the response rates. In addition, the response rates observed in the pilot (see Appendix C2) were used to decide on wave 1 stratification as explained in the next section.

6.5.6. Wave 1 sample

On the basis of the opt-out and response rates observed in the pilot test, the research team determined the number of observational units to be drawn in each stratification category for the raw sample.

As in the pilot test, the information provided for each observational unit included the names of the focal youth and the parent (if there was a parent), and for each of these individuals a phone number (where available) and address. The stratum to which each observation belongs was also indicated by a stratum label A to F (see Table 2). Strata information was not visible to interviewers.

6.5.7. Wave 1 opt-out process

As with the pilot study, FaHCSIA sent a primary approach letter and information brochure¹⁹ to each member of the observational unit before the beginning of the wave 1 interview period. The letter contained an invitation to participate in the survey and an explanation of how to opt out. Since wave 1 and associated opt-out and interviewing activities were carried out over the period of 5 months (July to December 2006), it was decided that the sample for wave 1 should be drawn in four stages, each stage subsequently representing a separate mail-out in the survey opt-out process. In order to minimise the time-span between the receipt of the primary approach letter and the first contact by RMR, the mail-outs were conducted approximately 2 weeks apart (14 July, 28 July, 11 August and 1 September 2006) – the first three mail-outs each containing 30 per cent of the total sample, and the last one containing 10 per cent. The opt-out process was then conducted separately for each mail-out stage.

¹⁹ The letter and information brochure can be viewed at <http://youthinfocus.anu.edu.au/survey.htm>

Overall, 13,568 young people born between 1 October 1987 and 31 March 1988 and their parent were drawn into the sample for wave 1 by Centrelink. As outlined in the previous section, the number of focal youth in each stratum were determined based on the pilot response and opt-out rates. The first five rows of Tables 3a and 3b contain detailed information on the number of focal youth and parents in each mail-out.

FaHCSIA maintained a 1800 number throughout the duration of the opt-out period, so that recipients of the primary approach letters could contact FaHCSIA to opt out or raise any queries about the survey. From the date the first mail-out sample was provided to RMR, a second 1800 number was set up, staffed by the market research firm, to liaise with respondents.

As explained in Section 3.3, the accuracy of the mailing addresses of potential respondents could not be easily ensured, especially for those who had last received Centrelink benefits a significant time ago. Therefore, the intended recipients of all primary approach letters that were returned to FaHCSIA marked 'return to sender' were treated as opt-outs and taken off the list of people whose contact details were provided to RMR.

RMR received a sample that contained 12,698 focal youth and 12,276 parents (for details, see Tables 3a and 3b). However, in some cases no phone number could be found for either member of the observational unit. In addition, a few people (11 focal youth and 11 parents) called FaHCSIA to opt out after their opt-out deadline and were consequently removed from the sample by RMR. As a result of this 'cleaning' process, the final clean sample downloaded into CATI for interviewing included 11,757 focal youth and 11,588 parents (for details, see Tables 3a and 3b).

Although, as explained above, all individuals whose primary approach letter was sent back as 'return to sender' were removed from the sample, a certain number of respondents, when approached by the market research firm, could not remember receiving the survey invitation and brochure and requested that these materials be re-sent prior to the interview.

For these cases, RMR compiled weekly lists that were provided to FaHCSIA, who then sent another copy of the primary approach letter and brochure to the participants' specified current addresses. Subsequently the responsibility for re-sending the letters and brochures was transferred to RMR. In the last week of the telephone interviewing, due to time constraints, the respondents who requested another copy of the survey materials were offered the YIF survey website address where the brochure could be viewed and the opportunity to receive a paper copy after the interview. If the respondents did not agree to this arrangement, the interview was terminated and they were classified as survey refusals.

Table 3a. Wave 1 sample and mail-out structure for focal youth

Stratification category	A	B	C	D	E	F	Total
Mail-out 1	910	1,650	490	390	398	233	4,071
Mail-out 2	910	1,650	490	390	398	233	4,071
Mail-out 3	910	1,649	490	390	398	232	4,069
Mail-out 4	303	550	163	130	133	78	1357
Total Centrelink sample (a)	3,033	5,499	1,633	1,300	1,327	776	13,568
<i>Opt-out and return-to-sender (b)</i>	259	320	96	91	61	43	870
<i>% of the Centrelink sample</i>	8.5%	5.8%	5.9%	7.0%	4.6%	5.5%	6.4%
Sample provided to RMR (a-b)	2,774	5,179	1,537	1,209	1,266	733	12,698
<i>No phone number/late opt-out (c)</i>	229	390	95	90	88	49	941
<i>% of the RMR sample</i>	8.3%	7.5%	6.2%	7.4%	7.0%	6.7%	7.4%
Clean sample (a-b-c)	2,545	4,789	1,442	1,119	1,178	684	11,757

Table 3b. Wave 1 sample and mail-out structure for parents

Stratification category	A	B	C	D	E	F	Total
Mail-out 1	899	1,633	484	387	395	231	4,029
Mail-out 2	901	1,634	488	388	397	233	4,041
Mail-out 3	893	1,629	484	387	391	231	4,015
Mail-out 4	300	544	160	127	132	76	1,339
Total Centrelink sample (a)	2,993	5,440	1,616	1,289	1,315	771	13,424
<i>Opt-out and return-to-sender (b)</i>	310	442	127	111	90	68	1,148
<i>% of the Centrelink sample</i>	10.4%	8.1%	7.9%	8.6%	6.8%	8.8%	8.6%
Sample provided to RMR (a-b)	2,683	4,998	1,489	1,178	1,225	703	12,276
<i>No phone number/late opt-out (c)</i>	167	294	63	65	68	31	688
<i>% of the RMR sample</i>	6.2%	5.9%	4.2%	5.5%	5.6%	4.4%	5.6%
Clean sample (a-b-c)	2,516	4,704	1,426	1,113	1,157	672	11,588

6.5.8. Wave 1 data gathering procedure

RMR made a substantial effort to contact each observational unit selected for the sample. This involved:

- Finding a telephone number through both automated and manual ‘white pages’ searches using the respondent’s last name;
- Finding a telephone number for individuals within an observational unit by

asking the other member of the observational unit.

RMR, in collaboration with the research team, also developed protocol for determining that observational units or individuals were to be deemed non-contactable:

- If no phone number was available for either member of the observational unit, and the 'white pages' search was unsuccessful, the unit was dropped from the clean sample. Similarly for units consisting only of a focal youth where no phone number could be found.
- If only one member of the observational unit had a phone number and the automated 'white pages' search for the other person was unsuccessful, the existing number was used to get into contact with the other member of the unit.
- Similarly, if only one member of the observational unit had a phone number and the number for the other person was found incorrect/invalid, then the valid number was used to get into contact with the other member of the unit.
- If both members of the unit had incorrect/invalid numbers, the automated 'white pages' search for both names was run at periodic intervals.

Some observational units provided in the sample to RMR were out-of-scope for other reasons. An observational unit was considered to be out-of-scope if the focal youth was out-of-scope.

Any individual, a focal youth or a parent, was considered out-of-scope if the individual was:

- Deceased;
- Overseas with no prospect of return before the end of the survey (August 2008);
- Institutionalised with no prospect of release before the end of the survey.

Contact details for individuals in institutions (for example, those confined to a correctional facility or medical treatment facility) were kept and these individuals were included in the sample. This should enable normal interviewing to resume if, and when, those individuals return to living in private dwellings. Similarly, while sample members overseas were not contacted, the contact details for individuals who moved overseas were maintained so that if they move back to Australia, they can be interviewed in subsequent waves.

RMR, in consultation with the research team, developed a protocol for converting contacts with individuals to successful interviews with the entire observational unit. This protocol included procedures for contacting people at different times and on different days of the week. If a contact resulted in the phone number being engaged, reaching an answering machine, or not answered, the procedure involved a callback at a different time on the next day. The individual was considered uncontactable after 10 consecutive callbacks that did not result in a contact.

Once any attempt was made to interview an individual in an observational unit, all

individuals in that observational unit were followed through until either:

- The individual was successfully interviewed;
- The individual clearly refused;
- A sufficient number of callbacks have been made at different times and on different days to satisfy the research team that all reasonable means to contact and interview an individual have been made.²⁰

6.5.10. Interview length and outcomes of self-completion questionnaire

In wave 1, all interviews were conducted over the telephone, using computer-assisted telephone interviewing (CATI). After completing the interview over the phone, the youth respondents were offered a choice of answering the self-completion questionnaire either on paper (to be sent by mail) or on-line at a secure website administered by RMR. Of all youth respondents, 73.4 per cent filled out self-completion questionnaires. Of those, 48 per cent were filled out on-line and 52 per cent on paper.

For wave 1 (first interview) telephone questionnaires, the average interview length was 30 minutes for a focal youth and 35 minutes for a parent. The time required for filling out the self-completion questionnaire, whether on paper or on-line, was estimated at 10 minutes.

6.5.11. Timeline

The fieldwork for wave 1 primarily occurred between 15 August 2006 and 16 December 2006. Self-completion questionnaires were accepted by RMR until 15 January 2007.

6.5.12. Payment to respondents

In order to test the effect of payment to respondents on response rates and characteristics of individual survey participants, the pilot for wave 1 was designed so that a randomly chosen half of respondents in each stratum would be offered a payment of \$15 by the market research firm, and the other half would not. The response rates were uniformly higher in the paid groups, especially among respondents who were heavily exposed to the income-support programs, and therefore the decision was made to offer the \$15 payment to all participants of the main wave.

The payments were mailed to respondents by RMR in the form of a cheque upon the completion of the telephone interview (parents) or upon the receipt of the self-completion questionnaire by the market research firm (focal youth).

²⁰ Among potential respondents who were declared uncontactable with call outcomes recorded as engaged, no reply, answering machine, or fax, 11.81 per cent of focal youth and 8.61 per cent of parents received less than 9 callbacks, and 4.21 per cent of youth and 5.45 per cent of parents received less than 5 callbacks.

6.5.13. Foreign language interviews

During the wave 1 fieldwork, if a language could be determined and a bilingual interviewer was available, interviews were attempted in languages other than English. Seventy-nine parent interviews were conducted in languages other than English, most frequently in Cantonese or Vietnamese. Written translations were prepared and used for section G of the parent questionnaire, which was more difficult to translate, and the questions on the locus of control were not asked at all, since the level of precision required for translation of these questions was much higher compared to other parts of the questionnaire. The rest of the questionnaire was translated by bilingual interviewers while they conducted the interview.

For the respondents who did not know English sufficiently well to participate and could not be interviewed by a bilingual interviewer, a separate code was created to record language problems as a reason for the interview termination. Among parents, 410 interviews (3.54 per cent of the clean sample) were terminated for this reason, and among focal youth, the respective number of terminated interviews was 51 (0.43 per cent of the clean sample).

6.5.14. Wave 1 contact outcomes

A summary of the outcomes of the wave 1 fieldwork is provided in Tables 4a and 4b.

Table 4a. Wave 1 focal youth outcomes

Stratification category	A	B	C	D	E	F	Total
Per cent of complete surveys across strata	25.2%	36.1%	12.9%	10.3%	9.9%	5.7%	
Completes	1,027	1,472	526	420	404	231	4,079
<i>% of clean strata sample</i>	40.4%	30.7%	36.5%	37.5%	34.2%	33.8%	34.7%
Self-completion questionnaire:	840	1,009	379	322	281	154	2,985
<i>online questionnaire</i>	509	433	209	173	163	84	1,571
<i>mail questionnaire</i>	331	576	170	149	118	70	1,414
Refusals	586	1,017	323	256	278	137	2,597
<i>% of clean strata sample</i>	23.0%	21.2%	22.4%	22.9%	23.6%	20.0%	22.1%
Unable to participate^a	103	185	61	30	45	24	448
<i>% of clean strata sample</i>	4.1%	3.9%	4.2%	2.7%	3.8%	3.5%	3.8%
Uncontactable^b	631	1,480	378	297	339	213	3,338
<i>% of clean strata sample</i>	24.8%	30.9%	26.2%	26.5%	28.8%	31.1%	28.4%
Unobtainable^c	198	635	154	116	113	79	1,295
<i>% of clean strata sample</i>	7.8%	13.3%	10.7%	10.4%	9.6%	11.6%	11.0%
TOTAL CLEAN SAMPLE	2,545	4,789	1,442	1,119	1,178	684	11,757

Notes:

a 'Unable to participate' category includes individuals who are out-of-scope (and observational units where the youth is out-of-scope), people who are unwell with no prospect of recovery until the end of the survey, are overseas or institutionalised with no prospect of return until the end of the survey, or do not speak English sufficiently well to participate.

b Individuals are declared 'uncontactable' if the number listed for them is a business number, if the named respondent is not at the number, or if, after a sufficient number of callbacks, the call outcomes on all callbacks are engaged, no reply, answering machine, fax or modem.

c Individuals are declared 'unobtainable' after a sufficient number of callbacks if the call outcomes on all callbacks are carrier messages like 'this number is disconnected', 'this number is uncontactable', or abnormal dialing tones that are not faxes or modems.

Table 4b. Wave 1 parent outcomes

Stratification category	A	B	C	D	E	F	Total
Completes	1,045	1,413	491	419	384	212	3,964
<i>% of clean strata sample</i>	41.5%	30.0%	34.4%	37.7%	33.2%	31.6%	34.2%
Refusals	741	1,281	428	350	376	212	3,388
<i>% of clean strata sample</i>	29.5%	27.2%	30.0%	31.5%	32.5%	31.6%	29.2%
Unable to participate^a	106	329	122	44	57	31	689
<i>% of clean strata sample</i>	4.2%	7.0%	8.6%	4.0%	4.9%	4.6%	6.0%
Uncontactable^b	366	899	221	178	211	125	2,000
<i>% of clean strata sample</i>	14.6%	19.1%	15.5%	16.0%	18.2%	18.6%	17.3%
Unobtainable^c	258	782	164	122	129	92	1,547
<i>% of clean strata sample</i>	10.3%	16.6%	11.5%	11.0%	11.2%	13.7%	13.4%
TOTAL CLEAN SAMPLE	2,516	4,704	1,426	1,113	1,157	672	11,588

Notes:

See Table 4a.

6.5.15. Data linkage

After completing the telephone survey, the respondents were asked for their consent to link their survey answers to the administrative data. The overwhelming majority of respondents, both in the pilot and the main wave, agreed to the linking. Table 5 below details percentages, by income-support history category, of respondents who consented to the data linkage.

Table 5. Respondents' consent to data linkage, per cent of total

Stratification category	A	B	C	D	E	F	Total
Pilot							
Focal youth	92.86	92.11	97.22	95.92	100.00	96.30	95.45
Parents	92.98	100.00	94.00	90.24	94.44	96.88	94.51
Main fieldwork							
Focal youth	97.47	95.45	95.44	96.19	96.77	96.10	96.20
Parents	93.49	91.15	91.85	93.08	94.53	92.45	92.46

6.5.16. Tracking and tracing

In order to minimise attrition between waves, both the pilot and the main wave respondents are tracked using a multiple-step follow up procedure. The timetable for the pilot keep-in-touch program was as follows:

Wave 1 pilot	March-April 2006
Mail-out of 'Thank you' card	July 17, 2006
Mail-out of Newsletter	December 11, 2006
Phone contact	October 2007
Wave 2 pilot	April-May 2008

The timetable for the main wave keep-in-touch program was as follows:

Wave 1 fieldwork	August-December 2006
Mail-out of 'Thank you' card	May-June 2007
Phone contact	September-October 2007
Mail-out of Newsletter	June-July 2008
Wave 2 fieldwork	August 2008-February 2009

6.6. Wave 2 fieldwork

6.6.1. Interviewer training

For the wave 2 fieldwork, RMR primarily used the same interviewers who were employed in the wave 1 fieldwork and therefore were familiar with the project. A training session for the interviewers was conducted on 28 August 2008 and was attended by Drs Robert Breunig and Anastasia Sartbayeva as well as the representative of the FaHCSIA Youth Bureau. An additional training session was held on 27 May 2009, before the beginning of the new entrant fieldwork.

6.6.2. Wave 2 pilots

Given the longitudinal nature of this project, the focal youth respondents of the wave 1 pilot were approached again for the continuing respondents pilot in wave 2. This pilot test was carried out between 21 April and 20 May 2008.

A pilot test for new entrants was carried out between 22 August and 30 September 2008. The sample for the pilot consisted of young adults randomly selected from the pilot birth cohort. From this sample, RMR excluded youth who completed the interview at wave 1 pilot (they were contacted as part of the continuing respondents pilot), youth who agreed to participate in wave 1 pilot but were not interviewed, youth who were deemed out of scope (overseas, deceased, institutionalised or unwell with no prospect of recovery until the end of the survey), and youth who refused to participate in the survey on such grounds as subject matter. The final new entrant sample consisted of 622 20-year olds who were either uncontactable at wave 1 pilot or gave a 'soft' refusal (lack of time, etc.) to participate.

6.6.3. Wave 2 pilot outcomes

Out of 242 18-year olds who participated in the wave 1 pilot, two were out of scope for wave 2 and therefore attempts were made to contact 240 respondents. The phone calls revealed 3 more respondents being out of scope for wave 2.

The continuing respondents pilot was run in two phases: phase one of the Computer Assisted Telephone Interviewing (CATI) was conducted between 21 April and 20 May 2008 and resulted in 157 completed surveys, making the response rate 66 per cent. This was lower than the anticipated rate of 75 per cent, likely due to the extension of the gap between the two waves from one to two years.

The second phase involved sending letters to both youth respondents and their parents, with the purpose of finding out the best contact number for the youth. Of the 72 letters mailed out, six resulted in interviews bringing the response rate to 68 per cent.

The pilot also involved testing of the self-completion questionnaire (SCQ), which the respondents could complete either by mail or on-line. The deadlines for the continuing respondents' SCQ receipt by Roy Morgan were 30 May 2008 for the on-line questionnaire and 2 June 2008 for the mail-out questionnaire. Of the 157 SCQs sent out after phase one, 69 per cent were returned. About 57 per cent of respondents chose the on-line option.

Of 622 respondents identified for contact in the new entrant wave 2 pilot, 595 were deemed to be in scope (available for the duration of the survey and without any language barrier); 135 interviews were collected, resulting in a response rate of 23 per cent. Since the market research firm did not implement any keep-in-touch strategy (refer to Section 6.5.16 for the description of the keep-in-touch program used for the continuing respondents) for this type of respondent, contact details had to be updated using a combination of 'white pages' searches and address and phone number updates provided by Centrelink.

6.6.4. Wave 2 sample

The continuing respondents sample consisted of 4,007 respondents. Of 4,079 focal youth who took part in Wave 1, 65 respondents had requested that the market research firm stop contacting them during the keep-in-touch program, and a further 7 respondents were deemed to be out of scope (3 were institutionalised, 2 were unwell with no prospect of recovery, and 2 were deceased). The contact details of respondents were updated through the keep-in-touch program described in Section 6.5.16, as well as through the update received from Centrelink.

The new entrant sample was based on the 11,757 youth respondents in the clean data sample for wave 1 (see Table 4a). From these, 5,761 respondents were excluded because they either (i) completed the wave 1 survey; (ii) gave a 'hard' refusal to participate; or (iii) were out-of-scope for wave 2 (deceased, or overseas, institutionalised or unwell with no prospect of recovery until the end of all waves of

the survey). The contact details for the remaining 5,996 young adults were loaded into the CATI system for the new entrant interviewing. In addition, the wave 2 new entrant sample included 583 young persons from the wave 1 raw sample who did not have a phone number on file at the time of wave 1 but for whom there was a contact number available in the address and phone number updates provided by Centrelink before wave 2. In total, the new entrant sample consisted of 6,579 individuals.

6.6.5. Wave 2 data gathering procedure

The general data gathering procedures were consistent with those implemented in wave 1 (refer to Section 6.5.8).

As a result of the keep-in-touch program and updated contact details provided by Centrelink prior to fieldwork, there were several contact numbers available for each participant in most cases, among them the young person's own number, numbers for alternate contacts, and the parent's number. Each available telephone number was attempted 3 times in total.

Calls to respondents took place between 5pm and 9pm on weeknights and between 11am and 6pm on weekends. If the telephone call was at an inconvenient time for the respondent, appointments were booked to contact the respondent at a time convenient for them, including day-time appointments as required. Respondents were called on both land-lines and mobile phones.

Due to the continuing respondents attrition rate being higher than the anticipated 25 per cent, it was decided to add an additional stage to the continuing respondents fieldwork. In this Stage 2 of the fieldwork, uncontactable youth were sent a letter with the purpose of finding out their best contact number(s), as well as the best time(s) to ring them. This resulted in achieving 45 additional interviews.

6.6.6. Interview length and outcomes of self-completion questionnaire

Telephone interviews for the 2,362 continuing respondents took, on average, just over 27 minutes. At the end of the interview, respondents were asked whether they would prefer to answer the self-completion survey on paper (to be sent by mail) or on-line at a secure website administered by RMR. More respondents (63 per cent) chose the on-line rather than the mail-out survey option. The overall response rate for the self-completion questionnaire was 71 per cent; 58 per cent among those who chose the paper option, and 78 per cent for those who chose the on-line option.

For the 1,261 new entrants, the average length of the telephone interview was 35 minutes. The majority of new entrant respondents (59 per cent) chose to complete the self-completion survey on-line as opposed to receiving a hardcopy version by mail. The overall response rate for the self-completion questionnaire among new entrants was 64 per cent; 52 per cent among those who chose the paper option, and 72 per cent for those who chose the on-line option.

6.6.7. Timeline

The main wave fieldwork for the continuing respondents was carried out between 28 August 2008 and 15 April 2009 and resulted in 2,362 completed phone interviews. The self-completion questionnaires were collected through the end of April 2009.

The new entrant fieldwork commenced on 27 May 2009 and ended on 4 August 2009. Self-completion questionnaires were collected until 5 September 2009.

6.6.8. Payment to respondents

As was the case during wave 1, continuing youth respondents were offered a 'thank you' payment of \$15 for participation in wave 2 of the survey. The payment was mailed out in the form of a cheque after RMR had received the self-completion questionnaire. However, after the conclusion of the continuing respondents fieldwork, RMR suggested increasing the payment to \$25 for those respondents who return the self-completion questionnaire, and \$15 for respondents who participate in the telephone interview only. These new payment arrangements were put in place starting with the new entrants survey.

6.6.9. Foreign language interviews

The interviews in wave 2 were conducted in English only. Among the continuing respondents, 7 interviews (less than 0.2 per cent of the in-scope sample) were terminated due to a language problem. For the new entrants, this number was slightly higher: 32 interviews were terminated due to a language problem (0.5 per cent of the new entrant sample).

6.6.10. Wave 2 contact outcomes

A summary of the outcomes of the wave 2 fieldwork for continuing respondents and new entrants is provided in Tables 6a and 6b.

Table 6a. Wave 2 continuing respondents outcomes

Stratification category	A	B	C	D	E	F	Total
<i>Percent of complete surveys across strata</i>	29.3%	32.0%	13.0%	10.8%	9.3%	5.5%	
Complete	692	757	308	256	219	130	2,362
<i>% of total loaded into CATI</i>	68.2%	52.4%	60.6%	61.7%	54.9%	57.5%	58.9%
Self-completion questionnaire:							
<i>online questionnaire</i>	515	521	211	194	152	84	1,677
<i>mail questionnaire</i>	376	329	153	139	114	61	1,172
	139	192	58	55	38	23	505
Out of scope	13	13	7	1	5	2	41
<i>% of total loaded into CATI</i>	1.3%	0.9%	1.4%	0.2%	1.3%	0.9%	1.0%
Uncontactable/unobtainable	15	36	8	10	14	5	88
<i>% of total loaded into CATI</i>	1.5%	2.5%	1.6%	2.4%	3.5%	2.2%	2.2%
Refusal/interview terminated	295	638	185	148	161	89	1,516
<i>% of total loaded into CATI</i>	29.1%	44.2%	36.4%	35.7%	40.4%	39.4%	37.8%
Total loaded into CATI at wave 2	1,015	1,444	508	415	399	226	4,007
<i>% of total wave 1 completes</i>	98.8%	98.1%	96.6%	98.8%	99.0%	97.8%	98.2%
Out of scope for wave 2, not loaded into CATI	12	28	18	5	4	5	72
<i>% of total wave 1 completes</i>	1.2%	1.9%	3.4%	1.2%	1.0%	2.2%	1.8%
Total wave 1 completes	1,027	1,472	526	420	403	231	4,079

Table 6b. Wave 2 new entrants outcomes

Stratification category	A	B	C	D	E	F	Total
<i>Percent of complete surveys across strata</i>	23.1%	41.6%	12.4%	9.2%	8.9%	4.8%	
Complete	291	525	156	116	112	61	1,261
<i>% of strata sample</i>	23.0%	17.9%	20.4%	19.8%	17.5%	15.7%	19.2%
Self-completion questionnaire:							
<i>online questionnaire</i>	210	306	110	74	67	38	805
<i>mail questionnaire</i>	163	174	83	48	43	21	532
	47	132	27	26	24	17	273
Unable to participate^a	52	78	36	32	15	13	226
<i>% of strata sample</i>	4.1%	2.7%	4.7%	5.5%	2.3%	3.3%	3.4%
Uncontactable^b	396	995	244	182	210	131	2,158
<i>% of strata sample</i>	31.3%	33.9%	31.9%	31.1%	32.9%	33.7%	32.8%
Unobtainable^c	226	782	157	121	128	90	1,504
<i>% of strata sample</i>	17.9%	26.6%	20.5%	20.7%	20.0%	23.1%	22.9%
Refused to participate	299	556	173	134	174	94	1,430
<i>% of strata sample</i>	23.7%	18.9%	22.6%	22.9%	27.2%	24.2%	21.7%
Total wave 2 sample	1,264	2,936	766	585	639	389	6,579

Notes:

a 'Unable to participate' category includes individuals who are out-of-scope (and observational units where the youth is out-of-scope), people who are unwell with no prospect of recovery until the end of the survey, are overseas or institutionalised with no prospect of return until the end of the survey, or do not speak English sufficiently well to participate.

b Individuals are declared 'uncontactable' if the number listed for them is a business number, if the named respondent is not at the number, or if, after a sufficient number of callbacks, the call outcomes on all callbacks are engaged, no reply, answering machine, fax or modem.

c Individuals are declared 'unobtainable' after a sufficient number of callbacks if the call outcomes on all callbacks are carrier messages like 'this number is disconnected', 'this number is uncontactable', or abnormal dialing tones that are not faxes or modems.

6.7. Future waves

At the time of writing, it is anticipated that wave 2 will be the last wave of the YIF survey. However, to insure for the possibility of future waves, the market research firm will maintain the contact details database of the survey participants and will implement a keep-in-touch program if wave 3 funding becomes available.

7. Survey dataset

7.1. Variable naming conventions

Each survey dataset begins with a set of variables that provide summary information on the respondent and the interview. Table 7 provides a list of these variables.

Table 7. Survey dataset introductory variable description

Variable name	Explanation
CentrelinkID	Random identification number assigned by Centrelink; postfixed by 0 for parent and 1 for focal youth
iiin	Permutation of the CentrelinkID; an integer variable that is the same for parent and focal youth belonging to the same observational unit
Surveyid	String version of the iiin variable
iisn	Survey number; identifies surveys of parents and focal youth, main vs. pilot stages of each survey wave, also identifies payment and non-payment respondent groups for wave 1 pilot
Economic	Stratification category (A to F) identifier
Pbmonth	(Parent survey only) Parent month of birth
Pbyear	(Parent survey only) Parent year of birth
Home_state	Home state of respondent
Home_post	Postcode of respondent
Ccount	Call count; total number of calls made to the respondent's telephone number
Avint	Time of the telephone interview
Totalltime	Total call time; time of all telephone calls made to the respondent
scr1a (youth), screx (parent)	Asks whether the respondent recalls receiving the survey invitation and information brochure
scr1, scr2	Asks whether the respondent is willing to conduct the telephone interview and/or arranges suitable appointment

The rest of the variables in the survey datasets are named according to the following naming scheme: wave/questionnaire identifier, section identifier, question number.

Wave/questionnaire identifier is represented by a single letter 'a' through 'z' and corresponds to:

a – parent telephone questionnaire, wave 1 pilot;

- b – youth telephone questionnaire, wave 1 pilot;
- c – youth self-completion questionnaire, wave 1 pilot;
- d – parent telephone questionnaire, wave 1;
- e – youth telephone questionnaire, wave 1;
- f – youth self-completion questionnaire, wave 1;
- g – youth continuing respondent telephone questionnaire, wave 2 pilot;
- h – youth continuing respondent self-completion questionnaire, wave 2 pilot;
- i – youth continuing respondent telephone questionnaire, wave 2;
- j – youth continuing respondent self-completion questionnaire, wave 2;
- k – youth new entrant telephone questionnaire, wave 2 pilot;
- l – youth new entrant self-completion questionnaire, wave 2 pilot;
- m – youth new entrant telephone questionnaire, wave 2;
- n – youth new entrant self-completion questionnaire, wave 2.

The wave/questionnaire identifier is followed by a section identifier which is a single letter 'a' through 'z'. The question number follows as it appears in questionnaire, i.e. '24a'. If the question is repeated several times (e.g., for different members of household) a '_#' is added to indicate which time it appears.

So for example, the question

J22: Do you only pay, only receive OR both pay AND receive any financial support for this child, excluding any money from the government?

in the youth telephone questionnaire for wave 1 would have a variable name ej22_1 when asked for the first child, ej22_2 when asked for the second child, etc.

The data file is rectangular. In other words, the maximum number of variables are created for each respondent and are coded missing for anyone of whom the question was not asked. For example, the maximum number of children the focal youth may have was set at 8, thus ej22_1 to ej22_8 exist for each observation even though the data would be largely missing for individuals with few or no children.

7.2. Labeling

The variables are labelled in the following way:

'Question #: text'

These labels have a limit of 80 characters including the question number, colon and spaces. If the question is more than 80 characters, it is abbreviated or shortened in a way that keeps the question identifiable. So for the above example, the label is

'ej22_1: do you pay, receive or both any financial support for this child?'

7.3. Survey datafiles characteristics

For each stage of the survey, RMR provided the ANU with two computer files, the first of which contains parent CATI survey responses and the second, focal youth CATI and self-completion questionnaire responses. Thus, after the completion of the pilot and main fieldwork for waves 1 and 2, the ANU has nine survey datafiles. Table 8 outlines the existing datafiles and the numbers of variables and records available in each. Note that for the focal youth who completed the telephone interview but not the self-completion questionnaire, the variables corresponding to the self-completion questionnaire in the focal youth datafiles contain missing data.

Table 8. Survey datafiles characteristics

Dataset	No. of variables	No. of records
Parent wave 1 pilot datafile	573	255
Focal youth wave 1 pilot datafile (both telephone and self-completion surveys)	699	242
Parent wave 1 datafile	590	3,964
Focal youth wave 1 datafile (both telephone and self-completion surveys)	785	4,079
Continuing respondents wave 2 pilot datafile	775	157
Continuing respondents wave 2 datafile	823	2,362
Continuing respondents wave 2 incomplete interviews datafile	823	1,645
New entrants wave 2 pilot datafile	598 (no SCQ data)	135
New entrants wave 2 datafile	853	1,261

Appendix A. Administrative dataset

Tables A1 and A2 summarise the total numbers of focal youth and primary parents in the administrative dataset (TDS2). These can be used as weights in the analysis of the YIF survey data.

The tables provide information by gender, month of birth (for the youth) and stratification category. In the first, trial version of TDS2 received by the ANU in October 2005, the stratification category variable was included in the dataset, while the later versions of TDS2 provided to the ANU did not contain the stratification variable. The ANU research team have re-created the stratification variable using the code provided by the Centrelink officer responsible for the drawing of the survey sample.

The stratification variable was generated using the September 2006 version of TDS2 for the youth in the main birth cohort and their primary parents according to the following procedure. First, a primary parent was identified for each focal youth using the criterion of the longest duration of care for the youth. If two or more parents had equal care duration, female parent was chosen over male (if more than one parent with equal duration of care were female, the primary parent was identified as the parent with the care period closest to the date of birth of the focal youth). Once all primary parents were identified, they were stratified according to their family income-support history (total duration of income-support receipt and the date of the first exposure). Finally, strata for primary parents were merged with youth records, each focal youth being assigned the same stratum as their selected primary parent. Youth with no recorded parent (2,227 cases), or those whose recorded parents did not have a TDS record (37 cases), were classified as category A.

In order to verify the stratification procedure, the stratification variable generated by the ANU was compared to the one included in the raw sample dataset drawn prior to wave 1 of the YIF survey, in July 2006. Out of 26,987 records of focal youth and primary parents in the raw sample file, we have successfully matched the strata for all but 161 cases. In almost all of these cases, the reason for the mismatch was the choice of a 'wrong' primary parent, due, perhaps, to the ANU not having access to the full set of Centrelink records. This, however, had almost no effect on the stratification of youth (with the exception of 8 cases), since both the 'wrong' and the 'right' primary parent had the same family income-support history data.

Table A1. Total numbers of focal youth in the administrative dataset (main birth cohort) by gender, month of birth and stratification category

Stratification category	Month of birth	Gender		Total
		Female	Male	
All categories	October 1987	10,300	11,225	21,525
	November 1987	9,924	10,315	20,239
	December 1987	10,138	10,911	21,049
	January 1988	10,451	10,884	21,335
	February 1988	10,147	10,428	20,575
	March 1988	11,344	11,759	23,103
	Total	62,304	65,522	127,826
Category A <i>No parental income-support history</i>	October 1987	4,308	4,699	9,007
	November 1987	4,095	4,288	8,383
	December 1987	4,091	4,428	8,519
	January 1988	4,277	4,450	8,727
	February 1988	4,147	4,270	8,417
	March 1988	4,659	4,755	9,414
	Category total	25,577	26,890	52,467
Category B <i>Heavy exposure to income-support programs (more than six total years on income support)</i>	October 1987	2,721	3,011	5,732
	November 1987	2,763	2,828	5,591
	December 1987	2,898	2,981	5,879
	January 1988	2,944	2,966	5,910
	February 1988	2,794	2,905	5,699
	March 1988	3,103	3,302	6,405
	Category total	17,223	17,993	35,216

Table A1. (continued) Total numbers of focal youth in the administrative dataset (main birth cohort) by gender, month of birth and stratification category

Stratification category	Month of birth	Gender		Total
		Female	Male	
Category C	October 1987	850	870	1,720
<i>First exposure to income-support system after 1998</i>	November 1987	786	787	1,573
	December 1987	794	923	1,717
	January 1988	877	914	1,791
	February 1988	840	810	1,650
	March 1988	968	927	1,895
	Category total	5,115	5,231	10,346
Category D	October 1987	868	945	1,813
<i>First exposure to income-support system between 1994 and 1998 and less than three total years on income support</i>	November 1987	871	872	1,743
	December 1987	883	932	1,815
	January 1988	843	982	1,825
	February 1988	848	871	1,719
	March 1988	930	1,037	1,967
	Category total	5,243	5,639	10,882
Category E	October 1987	1,075	1,161	2,236
<i>First exposure prior to 1994 and less than six total years on income support</i>	November 1987	970	1,052	2,022
	December 1987	951	1,090	2,041
	January 1988	994	1,017	2,011
	February 1988	1,043	1,060	2,103
	March 1988	1,121	1,178	2,299
	Category total	6,154	6,558	12,712
Category F	October 1987	478	539	1,017
<i>First exposure to income-support system between 1994 and 1998 and more than three but less than six total years on income support</i>	November 1987	439	488	927
	December 1987	521	557	1,078
	January 1988	516	555	1,071
	February 1988	475	512	987
	March 1988	563	560	1,123
	Category total	2,992	3,211	6,203

Table A2. Total numbers of primary parents in the administrative dataset (main birth cohort) by gender and stratification category

Stratification category	Gender			Total
	Not recorded	Female	Male	
Category A	6	47,863	1,756	49,625
Category B	7	33,643	1,078	34,728
Category C	1	9,601	599	10,201
Category D	3	10,515	205	10,723
Category E	4	12,401	152	12,557
Category F	3	5,856	264	6,123
Total	24	119,879	4,054	123,957

Appendix B. YIF questionnaires summary

B1. Parent questionnaire

Identification

- Date of birth, gender
- State and postcode
- Relationship to focal youth
- Marital status

Section A – Household information

- For each person living with: age, gender, relationship to respondent, identifier for the other natural parent of the focal youth

Section B – Personal information

- Country of birth (COB)
- If COB not Australia:
 - Date of first arrival in Australia
 - Language spoken in the family while growing up
 - English knowledge
- Aboriginal/Torres Strait Islander status
- Respondent's parents' occupations and family holidays at respondent's age 14

Section D – Education and employment

- Highest year of school completed and the type of school
- Highest educational degree attained
- Employment status and number of jobs
- Hours and earning in all jobs / main job
- Contract and salary arrangements
- Occupation
- Job search and main activity (if not employed)

Section E – Income from other sources

- Combined own and partner's income for 2005-2006 from wages/salary, own business, workers' compensation, interest or dividends, rental property etc.

Section F – Children and relationships

If respondent is a natural parent of focal youth:

- Total number of children
- Number of marriages/de facto relationships
- For each marriage/relationship:
 - Date of start / date when started living together
 - Reason and date of ending
 - Whether spouse/partner is the other natural parent
- For the other natural parent:
 - Country of birth
 - Education and employment
 - Amount of contact with respondent regarding the focal youth

If respondent is not a natural parent of focal youth, the following is asked about each natural parent:

- Relationship to respondent
- Country of birth
- Education and employment
- Date when stopped living with focal youth

Section G – Questions about Focal Youth’s education and youth

- Focal youth’s general performance in school and participation in extra-curricular activities
- Respondents’ participation in parent committees
- Remedial courses attended by focal youth
- Current enrolment status of focal youth
- Focal youth’s health: Asthma/ADHD/Depression diagnosis, disabilities, hospital visits, weight problems
- Date and reason focal youth moved out of parent’s/guardian’s house
- Type of accommodation focal youth lives in now
- Amount of respondent’s contact with focal youth and satisfaction with focal youth’s environment
- Quality of respondent’s relationship with focal youth
- Financial help to focal youth
- Teenager difficulties

Section H – General questions

- Attitudes on unemployment benefits
- Importance of own and parental education, ambition and job for success in life
- Locus of control questions

Section I – Health

- Overall rating of own health
- Smoking and drinking habits
- Asthma/depression/disability diagnosis; other health problems
- Height and weight

Section Y – Tracking information

B2. Youth wave 1 questionnaire

Section A – Identification

- Post code, state
- Date of birth
- Gender

Section B – Personal background

- Aboriginal/Torres Strait Islander status
- Country of birth (COB)
- If COB not Australia:
 - Date of first arrival in Australia

- Language spoken in the family while growing up

Section C – Current Household

- Marital/de facto status; gender and age of partner
- Living in large group/share accommodation
- Number of family members focal youth currently lives with and their age, gender and relationship to focal youth
- Living independently or with parents/guardians/parental figure; if lives independently – date and reason for moving out of the parents’ house

Section D – Childhood living arrangements

- Full/single parent/no parent family incidence
- Number of houses/apartments lived in

Section E – Family background

- Mother’s and father’s country of birth
- Whether lived with either or both parents at age 14
- Employment status of each parent (parent substitute) at focal youth’s age 14
- Parental education attainment (school and higher certificates/degrees)
- Quality of relationship with parents

Section F – School

- School enrolment status: year, type of school; overall number of schools attended
- Plans for completing school and continuing education
- Assessment of own performance at school: overall, mathematics, English
- Reasons for leaving school early
- Certificates attained as a result of studies

Section G – Post School student status and highest qualification

- Post-school studies: qualification, full-/part-time status, certificates
- Plans for completing higher degrees

Section H – Employment and earnings

- Employment status (past/present)
- Number of jobs
- Hours and earning in all jobs / main job
- Contract and salary arrangements
- Occupation and career
- First job history: age and hours

Section I – Job search and main activity

- Job search methods and length
- Main activity if not in school, employed, or looking for work

Section J – Family formation

- Date of start of current and first marriage/de facto relationship
- Number of marriages/de facto relationships
- Partner’s education and employment status
- Partner’s hours and earning in all jobs / main job
- Children: date of birth, living arrangements for the child, money paid/received for child support (other than government)

Section K – Housing and income

- Housing arrangements (renting/own)
- Rent/mortgage payment amounts
- Combined own and partner's income for 2005-2006 from wages/salary, own business, workers' compensation, investment etc.
- Financial help from parents

Section L – Health

- Overall health and whether it provides any limitations to work
- Effect of emotional problems on daily activities
- Emotional health in the past 4 weeks
- Presence of Asthma/ADHD/Depression diagnosis
- Smoking (yes/no)
- Height and weight

Section M – Attitudes

- Attitudes on unemployment benefits
- Importance of own and parental education, ambition and job for success in life

Section N – Tracking questions and SCQ

B3. Youth wave 1 self-completion questionnaire

Lifestyle and health

- Leisure and recreation activities
- Locus of control questions
- Physical activity
- Smoking and drinking habits
- Use of marijuana

Family and friends

- Quality of relationship with each parent/substitute

Education values

- Own, friends' and parents' attitudes to education

Life events and childhood

- Possible traumatic events that happened to focal youth
- Teenager problems and contact with authorities
- Relationship with parents during childhood
- Overall rating of childhood

B4. Continuing respondent (youth) wave 2 questionnaire

Section A – Identification

- Post code, state
- Date of birth
- Gender

Section C – Current Household

- Marital/de facto status; gender and age of partner
- Living in large group/share accommodation
- Number of family members focal youth currently lives with and their age, gender and relationship to focal youth
- Living independently or with parents/guardians/parental figure; if became independent since wave 1 – date and reason for moving out of the parents' house

Section D – Childhood living arrangements

- If moved out of parents' house since wave 1 – date of moving out

Section E – Family background

- Quality of relationship with parents

Section F – School

- School enrolment status: year, type of school; overall number of schools attended
- Plans for completing school and continuing education
- Assessment of own performance at school: overall, mathematics, English
- Reasons for leaving school early
- Certificates attained as a result of studies

Section G – Post School student status and highest qualification

- Post-school studies: qualification, full-/part-time status, certificates
- Plans for completing higher degrees

Section H – Employment and earnings

- Employment status (past/present)
- Number of jobs
- Hours and earning in all jobs / main job
- Contract and salary arrangements
- Occupation and career
- First job history: age and hours

Section I – Job search and main activity

- Job search methods and length
- Main activity if not in school, employed, or looking for work

Section J – Family formation

- Date of start of current and first marriage/de facto relationship
- Number of marriages/de facto relationships
- Partner's education and employment status
- Partner's hours and earning in all jobs / main job
- Children: date of birth, living arrangements for the child, money paid/received for child support (other than government)

Section K – Housing and income

- Housing arrangements (renting/own)
- Rent/mortgage payment amounts
- Combined own and partner's income for 2005-2006 from wages/salary, own business, workers' compensation, investment etc.
- Financial help from parents

Section L – Health

- Overall health and whether it provides any limitations to work
- Effect of emotional problems on daily activities
- Emotional health in the past 4 weeks
- Presence of Asthma/ADHD/Depression diagnosis
- Smoking (yes/no)
- Height and weight

Section M – Attitudes

- Most pressing current policies/issues
- Attitudes on unemployment benefits and immigration
- Importance of own and parental education, ambition and job for success in life
- Expectations of occupation and income in 10 years' time

Section N – Tracking questions and SCQ

B5. New entrant (youth) wave 2 questionnaire

Section A – Identification

- Post code, state
- Date of birth
- Gender

Section B – Personal information

- Country of birth (COB)
- If COB not Australia:
 - Date of first arrival in Australia
 - Language spoken in the family while growing up
 - English knowledge
- Aboriginal/Torres Strait Islander status
- Respondent's parents' occupations and family holidays at respondent's age 14

Section C – Current Household

- Marital/de facto status; gender and age of partner
- Living in large group/share accommodation
- Number of family members focal youth currently lives with and their age, gender and relationship to focal youth
- Living independently or with parents/guardians/parental figure; if lives independently – date and reason for moving out of the parents' house

Section D – Childhood living arrangements

- Full/single parent/no parent family incidence
- Number of houses/apartments lived in

Section E – Family background

- Mother's and father's country of birth
- Whether lived with either or both parents at age 14
- Employment status of each parent (parent substitute) at focal youth's age 14
- Parental education attainment (school and higher certificates/degrees)
- Quality of relationship with parents

Section F – School

- School enrolment status: year, type of school; overall number of schools attended
- Plans for completing school and continuing education
- Assessment of own performance at school: overall, mathematics, English
- Reasons for leaving school early
- Certificates attained as a result of studies

Section G – Post School student status and highest qualification

- Post-school studies: qualification, full-/part-time status, certificates
- Plans for completing higher degrees

Section H – Employment and earnings

- Employment status (past/present)
- Number of jobs
- Hours and earning in all jobs / main job
- Contract and salary arrangements
- Occupation and career
- First job history: age and hours

Section I – Job search and main activity

- Job search methods and length
- Main activity if not in school, employed, or looking for work

Section J – Family formation

- Date of start of current and first marriage/de facto relationship
- Number of marriages/de facto relationships
- Partner's education and employment status
- Partner's hours and earning in all jobs / main job
- Children: date of birth, living arrangements for the child, money paid/received for child support (other than government)

Section K – Housing and income

- Housing arrangements (renting/own)
- Rent/mortgage payment amounts
- Combined own and partner's income for 2005-2006 from wages/salary, own business, workers' compensation, investment etc.
- Financial help from parents

Section L – Health

- Overall health and whether it provides any limitations to work
- Effect of emotional problems on daily activities
- Emotional health in the past 4 weeks
- Presence of Asthma/ADHD/Depression diagnosis
- Smoking (yes/no)

- Height and weight

Section M – Attitudes

- Most pressing current policies/issues
- Attitudes on unemployment benefits and immigration
- Importance of own and parental education, ambition and job for success in life
- Expectations of occupation and income in 10 years' time

Section N – Tracking questions and SCQ

B6. Youth wave 2 self-completion questionnaire

Lifestyle and health

- Leisure and recreation activities
- Locus of control questions
- Physical activity
- Smoking and drinking habits
- Use of illicit drugs, including marijuana, ecstasy, cocaine, amphetamines/methamphetamines, and hallucinogens

Attitudes to risk

- Assessment of own risk-taking characteristics

Expectations

- Expectations of family formation in 10 years' time

Life events

- Possible significant events that happened to focal youth since turning 18 years of age and their effect on the youth's well-being

Appendix C. Additional information on wave 1 pilot

C1. Items tested in pilot

Survey respondent reaction to the primary approach letter and information leaflet.

For all of the survey instruments:

- Average interview length of the survey instruments;
- Routing procedures in the survey instruments so that all survey respondents were asked the questions appropriate to them;
- The response categories for scalar questions so that survey respondents were distributed across categories;
- Identify questions which generated high item non-response rates;
- Identify questions posing any other specific problem for survey respondents;
- Identify categorical questions which generated a large number of 'other' responses;
- Whether or not useful information was being obtained from open-ended questions;
- Any other survey instrument problems or issues.

Procedures for obtaining survey respondent consent to linking Centrelink administrative records with the survey data of the individual.

Procedures for contacting and interviewing individuals in observational units:

- Best methods to contact individuals;
- Proportion of invalid phone numbers;
- The result of all attempts to contact sample individuals and to track down individuals for whom phone number information was incorrect;
- Percentage of the provided sample that appeared to be uncontactable;
- The number of focal youth-parent interviews which were conducted on one and two different phone numbers;
- The number of observational units where complete responses were obtained from both individuals;
- The number of observational units where it was not possible to interview both individuals in the observational unit, the number of individuals who could not be interviewed, and the reason why they could not be interviewed;
- Compare the impact on response rates of offering a payment for participation.

The pilot study was also used to develop and test fieldwork manuals for the use of supervisors and interviewers.

Data coding and provision:

- Check that coding protocols do not produce any out-of-range codes;
- Test procedure and format for supplying data to the University.

C2. Pilot response rates

Table C2.1 below provides summary information on the strata-specific response rates for focal youth and parents interviewed during the wave 1 pilot. The table also outlines differences in response rates between the groups that were offered monetary compensation for participating in the survey and the groups that were not.

The response rates are calculated as a proportion of the clean sample. The clean sample is the original Centrelink sample minus those who chose to opt-out, those whose primary approach letter was returned to sender, observational units that consisted of parents without youth, and observational units for which no contact number could be found. Another procedure specific to the wave 1 pilot was that not everybody who agreed to participate in the survey was actually interviewed. The target numbers of interviews for both parents and youth (around 250) were reached before the entire pilot sample was exhausted, therefore, in order to more accurately assess response rates, RMR continued to approach individuals and record their refusal or consent to participate but did not actually interview them. The 'Completes' lines in Table C2.1 refer to the individuals who actually completed the interview as well as those who only agreed to be interviewed.

Table C2.1 Wave 1 pilot response rates

Stratification category	A	B	C	D	E	F	Total
Focal youth survey							
Paid respondents							
Clean sample size	97	97	99	102	97	101	593
Completes	39	32	36	45	45	39	236
Response rate	40.2%	33.0%	36.4%	44.1%	46.4%	38.6%	39.8%
Unpaid respondents							
Clean sample size	97	104	104	95	94	103	597
Completes	39	32	41	34	27	24	197
Response rate	40.2%	30.8%	39.4%	35.8%	28.7%	23.3%	33.0%
<i>Difference in response rate with respondent payment</i>	0.0%	2.2%	-3.1%	8.3%	17.7%	15.3%	6.8%
<i>Overall youth response rate</i>	40.2%	31.8%	37.9%	40.1%	37.7%	30.9%	36.4%
Parent survey							
Paid respondents							
Clean sample size	93	91	103	97	99	98	581
Completes	36	31	38	40	45	34	224
Response rate	38.7%	34.1%	36.9%	41.2%	45.5%	34.7%	38.6%
Unpaid respondents							
Clean sample size	89	100	102	100	96	102	589
Completes	34	21	40	26	31	38	190
Response rate	38.2%	21.0%	39.2%	26.0%	32.3%	37.3%	32.3%
<i>Difference in response rate with respondent payment</i>	0.5%	13.1%	-2.3%	15.2%	13.2%	-2.6%	6.3%
<i>Overall parent response rate</i>	38.5%	27.2%	38.0%	33.5%	39.0%	36.0%	35.4%

Table C2.2 summarizes strata-specific response rates for focal youth interviewed during the wave 2 continuing respondent and new entrant pilots. The table also shows additional interviews collected after the implementation of the stage 2 interviewing strategy for continuing respondents to ascertain the best contact number for the youth (refer to Section 6.6.3 for details).

Table C2.2 Wave 2 pilots response rates

Stratification category	A	B	C	D	E	F	Total
Continuing respondents pilot							
Stage One							
Sample size	55	38	34	48	36	27	238
Completes	39	24	23	34	22	15	157
<i>Response rate</i>	70.9%	63.2%	67.6%	70.8%	61.1%	55.6%	66.0%
Stage Two							
Additional completes	1	0	0	2	1	2	6
Total completes	40	24	23	36	23	17	163
<i>Final response rate</i>	72.7%	63.2%	67.6%	75.0%	63.9%	63.0%	68.5%
New entrants pilot							
Stage One							
Sample size	84	118	105	90	92	106	595
Completes	17	19	25	29	21	24	135
<i>Response rate</i>	20.2%	16.1%	23.8%	32.2%	22.8%	22.6%	22.7%

Bibliography

- An, Chong-Bum, Donna Ginther, Robert Haveman, and Barbara Wolfe, 1996. 'The 'Window Problem' in Studies of Children's Attainments: A Methodological Exploration', *Journal of the American Statistical Association*, 91, pp. 970-82.
- Australian Bureau of Statistics (ABS), 2002. *Births, Australia, 2002*
<http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/3301.02002?OpenDocument>
- , 2003. *Births, Australia, 2003*
<http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/3301.02003?OpenDocument>
- Centrelink, 2007. A guide to Australian Government Payments
[http://www.centrelink.gov.au/internet/internet.nsf/filestores/co029_0702/\\$file/co029_0702en.pdf](http://www.centrelink.gov.au/internet/internet.nsf/filestores/co029_0702/$file/co029_0702en.pdf)
- Gottschalk, Peter, 1996. 'Is the Correlation in Welfare Participation Across Generations Spurious?', *Journal of Public Economics*, 63, pp. 1-25.
- Haveman, Robert and Barbara Wolfe, 1995. 'The Determinants of Children's Attainments: A Review of Methods and Findings', *Journal of Economic Literature*, 33(4), December, pp. 1829-78.
- Haveman, Robert, Gary Sandelfur, Barbara Wolfe, and Andrea Voyer, 2001. 'Inequality of Family and Community Characteristics in Relation to Children's Attainments: A Review of Trends in Levels and Inequality in Characteristics and Attainments and Estimated Relationships between Them', University of Wisconsin, unpublished working paper.
- Israel, Mark and Michael Seeborg, 1998. 'The Impact of Youth Characteristics and Experiences on Transitions Out of Poverty', *Journal of Socio-Economics*, 27(6), pp. 753-76.
- McCoull, Frances and Jocelyn Pech, 2000. 'Trans-generational income support dependence in Australia: early evidence' in P Saunders (ed) *Reforming the Australian welfare state*, Australian Institute of Family Studies, Melbourne, pp. 91-111.
- Pech, Jocelyn and Frances McCoull, 1998. 'Transgenerational poverty and income support dependence in Australia: work in progress', *Social Security Journal*, 1998(2), pp. 167-82.

-----, 2000. 'Transgenerational welfare dependence: myths and realities', *Australian Social Policy*, 2000(1), pp. 43-68.

Solon, Gary, Mary Corcoran, Roger Gordon, and Deborah Laren, 1991. 'A Longitudinal Analysis of Sibling Correlations in Economic', *The Journal of Human Resources*, 26(3), pp. 509-34.

Wellman, Ray, 2005. 'Transgenerational Data Set: User documentation', version 1.0, 16/11/2005, Customer Data Services, Data Management Branch, FaHCSIA.