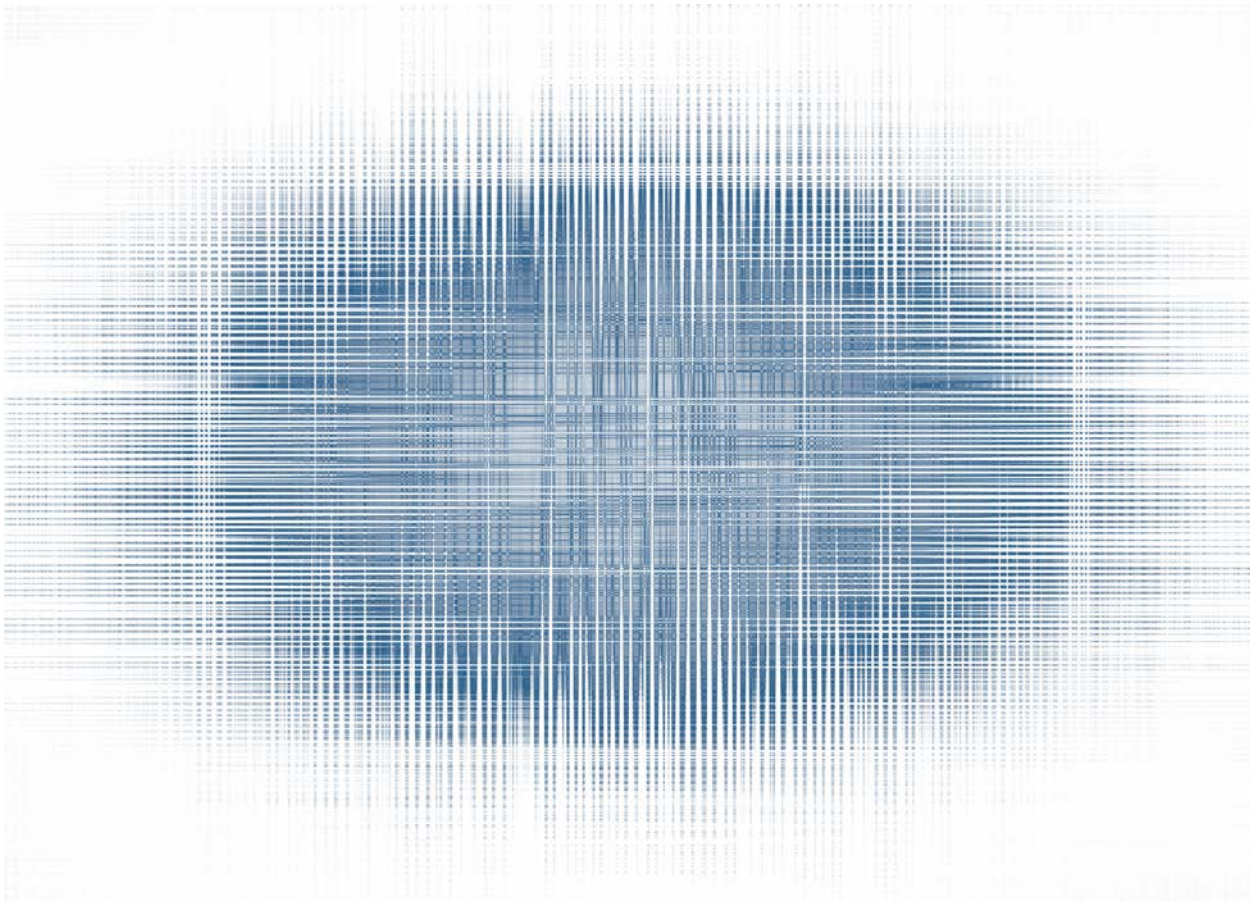
Australian National University
CENTRE FOR SOCIAL RESEARCH & METHODS



Consent to data linkage in a child cohort study, *Growing Up in Australia*: The Longitudinal Study of Australian Children

D Bandara, B Edwards, J Mohal and G Daraganova

CSRM WORKING PAPER
NO. 7/2019

Series note

The ANU Centre for Social Research & Methods (CSRM) was established in 2015 to provide national leadership in the study of Australian society. CSRM has a strategic focus on:

- development of social research methods
- analysis of social issues and policy
- training in social science methods
- providing access to social scientific data.

CSRM publications are produced to enable widespread discussion and comment, and are available for free download from the CSRM website (http://csrm.cass.anu.edu.au/research/publications).

CSRM is located within the Research School of Social Sciences in the College of Arts & Social Sciences at the Australian National University (ANU). The centre is a joint initiative between the Social Research Centre and the ANU. Its expertise includes quantitative, qualitative and experimental research methodologies; public opinion and behaviour measurement; survey design; data collection and analysis; data archiving and management; and professional education in social research methods.

As with all CSRM publications, the views expressed in this Working Paper are those of the authors and do not reflect any official CSRM position.

Professor Matthew Gray
Director, ANU Centre for Social Research & Methods
Research School of Social Sciences
College of Arts & Social Sciences
The Australian National University
November 2019

Working Paper No. 7/2019

ISSN 2209-1858
ISBN 978-1-925715-27-9

An electronic publication downloaded from http://csrm.cass.anu.edu.au/research/publications.

For a complete list of CSRM working papers, see http://csrm.cass.anu.edu.au/research/publications/working-papers.

ANU Centre for Social Research & Methods
Research School of Social Sciences
The Australian National University

Consent to data linkage in a child cohort study, *Growing Up in Australia*: The Longitudinal Study of Australian Children

D Bandara, B Edwards, J Mohal and G Daraganova

Dinusha Bandara is a Senior Manager of Data Management & Linkage, Longitudinal and Lifecourse Studies, at the Australian Institute of Family Studies.

Ben Edwards is an Associate Professor at the ANU Centre for Social Research & Methods.

Jatender Mohal is a Senior Manager of Data Management & Linkage, Longitudinal and Lifecourse Studies, at the Australian Institute of Family Studies.

Galina Daraganova is the Executive Manager of Longitudinal and Lifecourse Studies at the Australian Institute of Family Studies.

Abstract

In this paper, we provide new evidence on the factors associated with consent to data linkage in young people from a sample of 16–17-year-olds (born in 2004) participating in *Growing Up in Australia*: The Longitudinal Study of Australian Children. To our knowledge, this survey is the first time consent to data linkage has been studied in young people under 18 years of age. We extend the existing literature by examining economic record linkage, which is of particular concern in longitudinal surveys given the complexity of income support, benefits and pensions in many developed countries.

The findings show that young people's consent to data linkage is above 80%. The findings also shed light on significant demographic and psychosocial correlates of consent to data linkage, the influence of prior commitment to the survey by the household, the influence of previous consent to other forms of data linkage and the influence of family members' decisions on consent from the young person.

Acknowledgments

This paper makes use of data from *Growing Up in Australia*: The Longitudinal Study of Australian Children, which is conducted as a partnership between the Australian Government Department of Social Services, the Australian Institute of Family Studies and the Australian Bureau of Statistics, with advice provided by a consortium of leading researchers. Findings and views expressed in this publication are those of the individual authors only.

Acronyms

ACIR	Australian Childhood Immunisation Register
AEDC	Australian Early Development Census
ANU	Australian National University
CI	confidence interval
CSRM	Centre for Social Research & Methods
LSAC	The Longitudinal Study of Australian Children
MBS	Medicare Benefits Schedule
NAPLAN	National Assessment Program – Literacy and Numeracy
OR	odds ratio
PBS	Pharmaceutical Benefits Scheme
RPBS	Repatriation Pharmaceutical Benefits Scheme

Contents

Series note ii

Abstract iii

Acknowledgments iv

Acronyms iv

1 Introduction 1

1.1 Consent to data linkage 1

2 Methods 3

2.1 *Growing up in Australia*: The Longitudinal Study of Australian Children 3

2.2 Linked administrative data 3

2.3 Consent to data linkage 3

2.4 Factors influencing the study child's consent to link Centrelink income support administrative data 5

2.5 Statistical analyses 7

3 Results 10

3.1 Study child's consent to link income support administrative data 11

3.2 Role of prior data linkage consent 12

4 Discussion 13

Notes 15

References 16

Tables and figures

Table 1 National administrative data linked to LSAC 4

Table 2 LSAC consents by respondents and the two cohorts across waves 5

Figure 1 Framework of study child's consent to link income support administrative data 7

Table 3 Study child characteristics, parental characteristics, household characteristics and study commitment 8

Table 4 LSAC consent rates by administrative data, respondent and cohort 10

Table 5 Agreement between study child's consents at two consecutive waves 11

Table 6 Model results for association with study child's consent to link income support administrative data 12

1 Introduction

The next generation of longitudinal studies faces many of the challenges of other longitudinal surveys: declining response rates, increased survey costs to maintain survey response, and reduced budgets (Couper 2017). Longitudinal surveys, in particular, are expensive and require a high level of commitment from participants. The increasing availability of digital and administrative data provides new, rich data resources for researchers; an opportunity to reduce survey burden for participants; and a mechanism to accurately estimate the impact of survey attrition in subsequent waves (Calderwood & Lessof 2009, Davis-Kean et al. 2017). However, there is some evidence to suggest that the new generation of 'digital natives' (Prensky 2001) has increased expectations about their data privacy (Pew Research Centre 2017).

A small number of studies have investigated the factors associated with consent to data linkage in longitudinal surveys, particularly among young people. With the introduction of the new European Union General Data Protection Regulation, which focuses on explicit consent, more evidence is needed to understand factors associated with consent to data linkage for the next generation of longitudinal survey participants.

1.1 Consent to data linkage

There is some research examining demographic and psychosocial correlates of respondent consent to administrative data linkage in longitudinal studies, but very few studies examining consent in young people. These studies have found variation in the correlates, depending on the nature of the population and the type of consent (Sala et al. 2010, Baghal et al. 2014, Mostafa 2014, Baghal 2016, Mostafa & Wiggins 2017). For example, ethnic minorities are consistently reported as having lower levels of consent (Sala et al. 2010, Baghal et al. 2014, Mostafa 2014, Baghal 2016, Mostafa & Wiggins 2017), due to concerns about privacy and levels of trust (Sala et al. 2010, Baghal et al. 2014, Mostafa 2014). Some studies also report that participants with lower levels of education or socioeconomic status have lower consent rates (Baghal et al. 2014, Mostafa & Wiggins 2017). In household surveys, there has been some evidence of household 'contagion', with household members' decisions about whether or not to consent influencing others (e.g. Sala et al. [2010]).

We have identified only one study that has examined linkage consent in young people. Peycheva and colleagues (forthcoming) examined linkage consent rates when participants from the Next Steps longitudinal survey were aged 25–26 years. Consent to link was sought for nine administrative datasets. Using a data-driven approach with the rich data available from prior waves, the researchers identified many factors other than demographic and socioeconomic that were associated with consent. Many of these varied across different administrative data requests. For linkage to income support records, several novel factors were associated with higher likelihood of consent, including greater school engagement at age 13–14 years, answering sexual experience questions at 18 years, not in paid work at 17 years, seeking support when feeling down at 17 years, university enrolment by 25 years, not smoking at 25 years and residential stability at 25 years.

Correlates of consent rates vary, depending on the type of data linkage requested. The study by Peycheva et al. (forthcoming) identified few correlates associated with consent that are consistent across education, health and economic records.[1] Experimental and observational studies suggest that the highest rates of consent are for education records, followed by health records, and then income and economic records. In the only experimental study on this issue, Edwards and Biddle (forthcoming) randomly assigned participants to different groups and asked them to consent to different

types of data linkage. Participants' consent rates for education data linkage were 5% higher than those for health data linkage, and 20% higher than those for income support data linkage. Observational studies find the same pattern in household panel surveys (Sala et al. 2010), and among mothers in birth cohort studies (Mostafa 2014). Moreover, these patterns in consent to data linkage appear to be consistent across survey modes. For example, consent rates in the Next Steps survey, which uses a sequential, mixed-mode design, showed a similar variation, with the highest rates observed for education data linkage, followed by health and income/economic data (Peycheva et al. forthcoming).

Consistent with studies of response rates from different survey modes (Manfreda et al. 2008, Wengrzik et al. 2016), experimental and observational studies of data linkage consent find that consent rates are highest for face-to-face interviews. The only experimental study examining the effects of survey mode on consent rates found that consent rates were 30% lower for web surveys than for face-to-face surveys (Jäckle et al. 2018). Similarly, data linkage consent rates for the Next Steps survey were 30% lower on the web than for telephone or face-to-face surveys for income/pensions, health and education data (Peycheva et al. forthcoming).

Economic record linkage is of particular concern in longitudinal surveys because of the complexity of income support, benefits and pensions in many developed countries; the central nature of economic information to many longitudinal studies; and the fact that studies have shown that data linkage consent is lowest for economic records (Sala et al. 2010, Mostafa 2014, Edwards & Biddle forthcoming).

In this paper, for the first time, we examine consent rates to different types of administrative data linkage, and particularly factors associated with data linkage consent among young people (aged 16–17 years, born in 2004) participating in *Growing Up in Australia*: The Longitudinal Study of Australian Children. Consent was requested during face-to-face interviews, which was a novel approach to obtaining consent to income support data linkage from young people under 18 years of age. We consider demographic and psychosocial correlates of consent to data linkage. We also consider design features common to many longitudinal studies, including the influence of prior commitment to the study by the household, previous consent to other forms of data linkage, and the influence of family members' decisions on consent from the young person.

2 Methods

2.1 *Growing up in Australia*: The Longitudinal Study of Australian Children

The Longitudinal Study of Australian Children (LSAC) is Australia's first nationally representative (Edwards 2014) longitudinal study of child development. It was designed as a stratified random sample of children from all Australian states and territories (excluding the most remote areas). The sample was selected from the Medicare Australia (formerly Health Insurance Commission) database, the most comprehensive database of the Australian population. In 2004, LSAC recruited 5107 children aged 0–1 year (the baby or B cohort) and 4983 children who were 4–5 years old (the kindergarten or K cohort), and their families across all states and territories of Australia. Information on these participants has been collected every 2 years from multiple respondents, including resident and nonresident parents, teachers and carers, and by direct child assessment and self-report. Detailed descriptions of the study design and procedures can be found in LSAC technical papers (Soloff et al. 2005, Gray & Smart 2009).

2.2 Linked administrative data

Over the years, LSAC data have been linked to national administrative data from the Medicare Benefits Schedule (MBS), the Pharmaceutical Benefits Scheme (PBS), the Repatriation Pharmaceutical Benefits Scheme (RPBS), the Australian Childhood Immunisation Register (ACIR), the National Assessment Program – Literacy and Numeracy (NAPLAN), the Australian Early Development Census (AEDC), the Australian Curriculum Assessment and Reporting Authority (ACARA, also known as MySchool), and Centrelink. These sources of administrative data are described in Table 1.

2.3 Consent to data linkage

Table 2 provides summary information on LSAC consents collected by respondents across waves. In LSAC, a wave is defined as a point in time when researchers collect data about children's development within the current economic, social and cultural environment.

The study data were collected from multiple respondents. 'Study child' is the cohort child and the main respondent of the study. 'Parent 1' is the parent who knows the study child best; in most cases, this is the child's biological mother. 'Parent 2' is Parent 1's partner or another adult in the home with a parental relationship to the study child; in most cases, this is the biological father, but stepfathers are also common. Parent 1 (and/or Parent 2) might change between waves, and any new parent (new Parent 1) may join in subsequent waves.

At wave 1, Parent 1 consented to data linkage on behalf of the study child for the MBS, PBS and ACIR. Parents also consented to having their address details tracked. They had to complete the consent form and sign it in the presence of a witness. Incomplete forms resulted in incomplete consent. New parents who joined in subsequent waves also consented on behalf of the study child.

At wave 3 data collection, Parent 1 of the K cohort children was asked to provide consent for allowing access to their study child's NAPLAN data. For those who did not provide consent at wave 3, consent was obtained at wave 4. At wave 4 data collection, Parent 1 of the B cohort children was asked to consent to data linkage for NAPLAN and the AEDC on behalf of the study child. In subsequent waves, a new Parent 1 who joined the study also consented.

In wave 6, study children in the K cohort at the age of 14–15 years were asked to consent for the

Table 1 National administrative data linked to LSAC

Area	Name	Description
Medical	Medical Benefits Schedule (MBS)	Contains information on processed claims, patients and service providers for services that qualify for a benefit under the *Health Insurance Act 1973* (AIHW 2018a, MBS)
	Pharmaceutical Benefits Scheme (PBS)	Contains information on PBS scripts, payments, patients, prescribers and dispensing pharmacies for prescription medicines that qualify for a benefit under the *National Health Act 1953* (AIHW 2018b, PBS)
	Repatriation Pharmaceutical Benefits Scheme	Structured like the PBS, contains information on prescription medicines prescribed to Department of Veterans' Affairs (DVA) beneficiaries, including eligible veterans, war widows/widowers and their dependants, under the *Veterans' Entitlements Act 1986* (DVA[a])
	Australian Childhood Immunisation Register (ACIR)	Includes records of vaccinations given to children under 7 years of age who live in Australia and have been enrolled in Medicare since 1996 (ACIR[b])
Educational	National Assessment Program – Literacy and Numeracy (NAPLAN)	Includes educational achievement test records in literacy and numeracy from an annual assessment for students enrolled in years 3, 5, 7 and 9 since 2008 (National Assessment Program[c])
	Australian Early Development Census (AEDC)	Includes school readiness measures of early childhood development at the time children commence their first year of full-time school, every 3 years since 2009 (AEDC[d])
	Australian Curriculum Assessment and Reporting Authority (ACARA)[e]	Contains information about schools and the outcomes of schooling, as required by the Council of Australian Governments Education Council (ACARA[f]). ACARA is responsible for collating NAPLAN data received from Australian schools, collecting school characteristics and managing the MySchool website
Income support	Centrelink	Includes Centrelink income support payment records for services provided at times of major change for seniors, jobseekers, students and trainees, families, carers, parents, people with disability, Indigenous Australians, and people from culturally and linguistically diverse backgrounds (DHS[g])

a www.dva.gov.au/factsheet-hsv92-repatriation-pharmaceutical-benefits-scheme

b https://data.gov.au/dataset/australian-childhood-immunisation-register

c www.nap.edu.au

d www.aedc.gov.au

e ACARA is also known as MySchool. This linkage does not require consent.

f www.acara.edu.au/reporting/my-school-website

g Australian Government Department of Human Services (www.humanservices.gov.au/organisations/about-us/statistical-information-and-data)

first time to data linkage for the MBS, PBS and RPBS. In waves 5 and 6, a new Parent 1 who joined the study also consented.

In wave 7, Parent 1 and Parent 2 themselves consented to their data linkage for the MBS, PBS and RPBS. Consent to access NAPLAN data in wave 7 for the K cohort was not required because study children were beyond year 9 level at school (i.e. when NAPLAN does not apply). However, at this wave, consent to link income support administrative data was obtained from Parent 1 and Parent 2 of K cohort children. The K cohort study child aged 16–17 years also provided consent to link income support administrative data for the first time.

Table 2 **LSAC consents by respondents and the two cohorts across waves**

Wave	Respondent	Consent for	B cohort	K cohort
1	Parent 1	Study child	MBS, PBS, ACIR	MBS, PBS, ACIR
2	New Parent 1	Study child	MBS, PBS, ACIR	MBS, PBS, ACIR
3	New Parent 1	Study child	MBS, PBS, ACIR	MBS, PBS, ACIR
	Parent 1	Study child	No linkage	NAPLAN
4	New Parent 1	Study child	MBS, PBS, ACIR	MBS, PBS, ACIR
	Parent 1	Study child	NAPLAN, AEDC	NAPLAN
5	New Parent 1	Study child	MBS, PBS, ACIR, NAPLAN	MBS, PBS, ACIR, NAPLAN
6	New Parent 1	Study child	MBS, PBS, ACIR, NAPLAN	MBS, PBS, ACIR, NAPLAN
	Study child	Themselves	No linkage	MBS, PBS
7	New Parent 1	Study child	MBS, PBS, ACIR, NAPLAN	MBS, PBS, ACIR
	Study child	Themselves	No linkage	CLNK
	Parent 1/Parent 2	Themselves	MBS, PBS, RPBS	MBS, PBS, RPBS, CLNK

ACIR = Australian Childhood Immunisation Register; AEDC = Australian Early Development Census; CLNK = Centrelink income support administrative data; MBS = Medicare Benefits Schedule; NAPLAN = National Assessment Program – Literacy and Numeracy; PBS = Pharmaceutical Benefits Scheme

2.4 Factors influencing the study child's consent to link Centrelink income support administrative data

We developed a framework of consent (Figure 1) to explore the factors that influence consent to link to income support administrative data among study children aged 16–17 years, mindful that the framework should have the capacity to include novel variables (following Peycheva et al. forthcoming). This framework considered the interplay between the study child, parental characteristics, the household environment and study commitment, and the study child's consent.

The study child's birthweight, gender, birth order, Aboriginal/Torres Strait Islander status, any medical conditions/disabilities expected to last 6 months or longer, conduct problems, academic problems, and their previous consent to link MBS/PBS data were considered within the framework to be included in the analyses.

The limited prior research shows that school engagement is important for linkage consent to income support data (Peycheva et al. forthcoming). Therefore, we include a measure of conduct problems and academic problems to capture engagement in school (see Table 3). Using parents' responses to the 'Conduct Problems' subscale of the Strengths and Difficulties Questionnaire (Goodman 2001), we classified children into two categories: 'scores less than 4' and 'scores of at least 4'. Children were considered to have higher symptoms of conduct problems if they had scores of at least 4 (AMHOCN 2005). Using the rating of 'Overall School Achievement' from the Teacher Academic Rating Scale (used in the Early Childhood Longitudinal Study, Kindergarten cohort of 1988–99), we classified children into two categories: 'those having academic problems' and 'those not having academic problems'. Children were considered to have academic problems if they were rated as 'below average' or 'far below average', and as not having academic problems if they were rated as 'far above average', 'above average' or 'average' by their teacher.

In parental characteristics (see Figure 1 and Table 3 for full list), we included information from both Parent 1 and Parent 2, where information was available. Unless stated otherwise, characteristics from both parents were combined

to create three groups: 'common to both parents', 'common to only one parent' and 'not common to both parents'.

Other studies have reported lower rates of consent from people from diverse ethnic backgrounds (parent country of birth, language other than English, child Indigenous status) (e.g. Baghal et al. [2014]). Parents' levels of time pressure (weekly work hours from all jobs), stress (stressful life events, psychological distress) and substance use (current smoking status, alcohol consumption) could influence the degree to which they approach the interviewer in a hostile or suspicious manner and influence their children. We categorised alcohol consumption into 'high' and 'low' using information on regular short-term risky drinking – that is, whether the parent has five or more (for women) or seven or more (for men) drinks in a single sitting, on at least two or three occasions per month. Using the National Longitudinal Survey of Children and Youth (Statistics Canada 2000), we explored how often parents and primary caregivers engaged in a range of behaviours that demonstrate consistent parental discipline. We then categorised the data into 'scores less than 3.5', which indicates a low level of consistent discipline or inconsistent discipline, and 'scores of at least 3.5', which indicates consistent parental discipline.

Using responses to the Argumentative Relationship Scale (an adoptive measure of the co-parenting scale; Ahrons 1981), we placed respondents into two categories: those above the median in argumentativeness and those below the median. We then explored argumentativeness for both parents and classified their relationship into one of four categories: 'both parents reported low (i.e. below the median) argumentativeness', 'one parent reported high (i.e. above the median) argumentativeness', 'both parents reported high argumentativeness', and 'ineligible to answer the argumentative scale – lone-parent families'. Using the Kessler K6 Screening Scale (Kessler et al. 2003), we divided participants into two categories of psychological distress: 'those scoring at least 14' (who were considered to have psychological distress) and 'those scoring under 14' (who were not considered to have psychological distress). Using an adaptation of the Stressful Life Events scale from the Path Through Life Study (CMHR 2005), the total number of stressful life events was classified into 'less than four stressful events' and 'four or more stressful events'. The question was asked of Parent 1 but relates to both parents.

Although the inclusion of a measure of parenting style (parental consistency) as well as a measure of parental discord (parental argumentativeness) was exploratory, young people may be directly influenced by their interaction with their parents and through observing parents, and this may well translate into higher consent rates at interview. Parental consent to link MBS/PBS data was included, to capture parental influence (e.g. Sala et al. 2010).

The study child's household characteristics included lone-parent household, siblings in the household, equivalised household income quintiles, Socio-Economic Indexes for Areas (SEIFA) index of advantage/disadvantage, government benefits as family's main income, financial stress, remote residence, housing tenure and household mobility (whether study child has moved house in the previous 2 years). Household income was classified into quintiles of equivalised household income. The SEIFA Index of Relative Socio-economic Advantage and Disadvantage is an area-level socioeconomic measure (ABS 2008), which was classified into 'lowest 25%', 'middle 50%' and 'highest 25%'.

In terms of household characteristics, we include an indicator for receipt of government benefits, as some research has indicated that having a prior relationship with the agency may lead to higher rates of consent (Sakshaug et al. 2012, Mostafa 2014, Edwards & Biddle forthcoming). Having a family member with a disability or medical condition could also mean that the household is more likely to receive a government benefit of some form (Disability Support Pension, Carer Payment or Carer Allowance).

Using the financial hardship scale (Bray 2001), financial stress options were grouped into 'no stressful events', 'one event' and 'two or more events'. Housing factors (mobility, tenure, neighbourhood socioeconomic disadvantage and remoteness) have been found to influence consent (e.g. Peycheva et al. forthcoming).

As with previous research in this area, study commitment or loyalty is considered to be a key driver of future consent (Mostafa 2014). We used 'study commitment' as a covariate, which was defined using participation and nonresponse data from current and previous waves (see Figure 1). This variable had four categories: 'responded to all waves', 'marginally attached and responding', 'refusal in the last wave' and 'noncontact in the last wave' (Bandara et al. forthcoming). In the analyses, only the first two categories were included, because the subpopulation who were not available to consent at wave 7 was the same as those who refused or were noncontact in the last wave.

Table 3 presents the distribution of study child characteristics (*N* = 3089), parental characteristics, household characteristics and study commitment.

2.5 Statistical analyses

First, we used descriptive statistics to understand consent rates to different types of administrative data across waves, to provide a context for young people's data linkage consent to income support administrative data.

Second, to gain a preliminary understanding of the degree of consistency of data linkage consent over time, agreement between the study child's consent to link MBS and PBS data at wave 6 with consent to link income support administrative data at wave 7 was described with a 95% confidence interval (CI). Analysis only included children who participated in both waves (waves 6 and 7) and who provided consent information.

Third, in light of the framework (see Figure 1 and Table 3), univariate logistic regressions (Peters 2008) explored the association between each of the study child's characteristics (excluding previous consent to link MBS/PBS data), parental characteristics, household environment and study

Figure 1 **Framework of study child's consent to link income support administrative data**

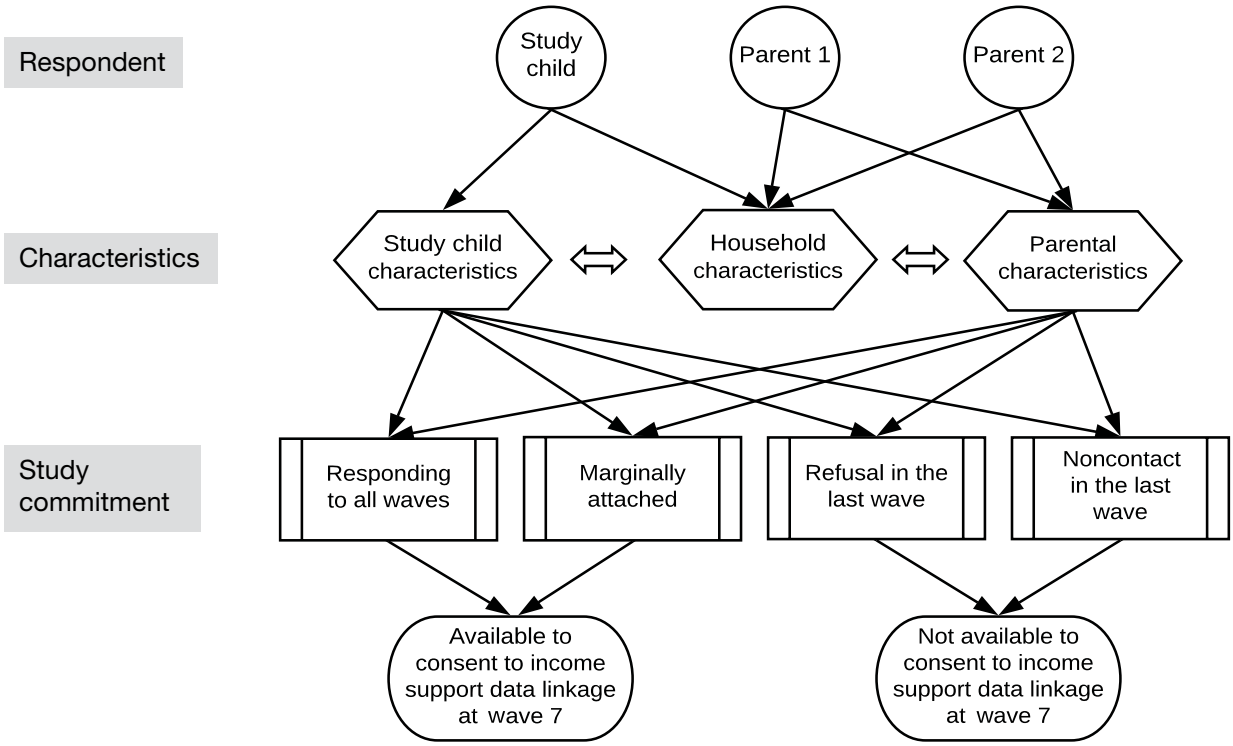


Table 3 Study child characteristics, parental characteristics, household characteristics and study commitment

Factor	Categories
Birthweight	Low birthweight (<2500 g) = 5.8%; appropriate birthweight (2500–4000 g) = 80.5%; high birthweight (>4000 g) = 12.7%; not reported = 1%
Gender	Male = 51%; female = 49%
Birth order	First-born child = 43.5%; second or subsequent child = 56.5%
Aboriginal or Torres Strait Islander status	Not Aboriginal/Torres Strait Islander = 97.7%; Aboriginal/Torres Strait Islander = 2.3%; not reported = 0.1%
Medical condition/ disability[a]	No medical condition/disability = 93.7%; medical condition/disability = 6.3%
Conduct problems	Scores of <4 = 93.7%; scores of ≥4 (higher symptoms of conduct problems) = 6.3%
Academic problems	Not having academic problems = 81.6%; having academic problems = 13.5%; not reported = 4.9%
Previous medical consent[b]	Did not consent = 10.8%; consented = 85.4%; did not participate at the time of consent = 3.9%
Parental country of birth	Both parents born in Australia/New Zealand = 61.5%; only one parent born in Australia/New Zealand = 26.3%; both parents born outside Australia/New Zealand = 12.1%; not reported = 0.03%
Parental main language spoken at home	Both parents spoke English = 81.1%; only one parent spoke a language other than English = 8.7%; both parents spoke a language other than English = 9.8%; not reported = 0.4%
Parental education	Year 12 or lower = 12.5%; advanced diploma/certificate/other = 40.5%; bachelor degree and above = 47%
Parental working hours	Both parents worked ≥35 hours = 36.4%; only one parent worked ≥35 hours = 54.4%; both parents worked <35 hours = 4.9%; not reported = 4.3%
Parental medical condition/disability[a]	Both parents with no medical condition/disability = 75.6%; one parent with medical condition/disability = 18.2%; both parents with medical condition/disability = 2.7%; not reported = 3.5%
Parental smoking status	Both parents do not smoke = 71.6%; only one parent currently smokes = 19.9%; both parents currently smoke = 5.3%; not reported = 3.1%
Parental alcohol consumption	No-one consumes high levels = 70.1%; only one parent consumes high levels = 17.8%; both parents consume high levels = 4.4%; not reported = 7.8%
Parental consistency	Both parents reported high consistency = 55.6%; one parent reported high consistency = 32.4%; both parents reported low consistency = 9.1%; not reported = 2.9%
Parental argumentativeness	Both parents reported low argumentativeness = 34.4%; one parent reported high argumentativeness = 25.9%; both parents reported high argumentativeness = 19.7%; lone-parent families = 19.1%; not reported = 0.9%
Parental psychological distress	Both parents with low level of psychological distress = 72.6%; one parent with high level of psychological distress = 22.8%; both parents with high level of psychological distress = 2.2%; not reported = 2.4%
Parental stressful life events	Less than 4 stressful events = 68.6%; 4 or more stressful events = 31%; not reported = 0.5%
Parental consent[c]	Both parents did not consent = 5.5%; one parent consented = 52.7%; both parents consented = 32%; parent(s) did not participate at the time of consent = 9.8%
Siblings in household	None = 16.7%; 1 sibling = 45.1%; 2 siblings = 26.6%; 3 or more siblings = 11.6%

continued

Table 3 ***continued***

Factor	Categories
Household income	1st (lowest) quintile = 20.3%; 2nd quintile = 20.3%; 3rd quintile = 19.8%; 4th quintile = 19.9%; 5th (highest) quintile = 19.7%
SEIFA index of advantage/ disadvantage	Lowest 25% = 25.6%; middle 50% = 50.5%; highest 25% = 23.9%
Government benefits[d]	No = 93.2%; yes = 6.7%; not reported = 0.2%
Financial stress	No stressful events = 85%; 1 event = 9%; 2 or more events = 6%
Remote residence	Moderate to highly accessible area = 96%; remote or very remote area = 3.6%; not reported = 0.4%
Housing tenure	Owner without a mortgage = 20.9%; owner with a mortgage = 58.7%; renter – private landlord = 14.2%; renter – state/territory housing authority = 2.2%; other landlord/other tenure type = 4.1%
Household mobility[e]	No = 83.6%; yes = 16.3%; not reported = 0.03%
Study commitment	Responded to all waves = 90.4%; marginally attached and responding = 9.6%

SEIFA = Socio-Economic Indexes for Areas

a Expected to last 6 months or longer

b Consent to link Medicare Benefits Schedule/Pharmaceutical Benefits Scheme data

c Consent to link income support administrative data

d Family's main income is government benefits.

e Whether study child has moved house in the previous 2 years. Information was sought from wave 7 K cohort data collection; if this was not available, information from the previous wave was used to capture appropriate previous longitudinal information.

Note: *N* = 3089.

commitment and the study child's consent to link income support administrative data.

Based on these analyses and significant findings, a mutually adjusted multivariable logistic regression model was developed. The results presented from the model include the Wald chi-square statistic, the *P* value of Wald chi-square statistic, adjusted odds ratios (OR) and 95% CI values. Where appropriate, 'not reported' categories within factors were included to capture the full possible analytical sample (*n* = 3087; excluding partial missing data: *n* = 2).

A full model was also developed that included all factors from initial analyses. However, because it did not improve the model fit, results are not presented.

Lastly, to understand the role of prior data linkage consent, a separate mutually adjusted model examined the study child's previous consent to link MBS/PBS data (at wave 6) with their consent to link income support administrative data. The model included all the factors from the last model, but the sample was restricted to children with available consent information from both wave 6 for MBS/PBS data linkage and wave 7 for income support administrative data linkage (*n* = 2967; excluding partial missing data: *n* = 2). The ORs and CIs are presented.

Analyses were conducted using SAS software (version 9.4, SAS Institute, Cary, NC, USA). A two-sided *P* value of <0.05 was considered statistically significant.

3 Results

Table 4 presents the LSAC consent rates by respondents and both cohorts across waves. Rates were calculated based on those who participated at the wave when the consent was sought.

Consent to link MBS, PBS and ACIR administrative data to LSAC was provided by 93% of Parent 1 of the B and K cohorts. Nearly 90% of B cohort parents provided consent to link AEDC and NAPLAN data, and 95% of K cohort parents provided consent to link NAPLAN data. Parent consent to link their MBS data was provided by 82% and 89% of Parent 1 of the B and K cohorts, respectively. For PBS/RPBS data, the figures were 81% and 88% of Parent 1 of the B and K cohorts, respectively. These rates were lower for Parent 2: 70% and 72% for MBS, and 68% and 69% for PBS/RPBS. K cohort Parent 1 had an 85% consent rate to link to income support administrative data, compared with

Table 4 **LSAC consent rates by administrative data, respondent and cohort**

			B cohort		K cohort	
Wave	Data	Respondent	*N*	Consent rate, % (95% CI)	*N*	Consent rate, % (95% CI)
Parent reporting on behalf of study child						
1	MBS	Parent 1	5107	93.4 (92.7–94.1)	4983	93.4 (92.7–94.1)
1	PBS	Parent 1	5107	93.3 (92.6–94.0)	4983	93.4 (92.7–94.1)
1	ACIR	Parent 1	5107	93.5 (92.8–94.2)	4983	92.3 (91.6–93.1)
4	AEDC, NAPLAN	Parent 1	4242	89.9 (89.0–90.8)	na	na
3/4	NAPLAN	Parent 1	na	na	4431	95.4 (94.8–96.0)
Parent reporting for themselves						
7	MBS	Parent 1	3286	82.0 (80.7–83.3)	3002	89.0 (87.9–90.2)
7	PBS/RPBS	Parent 1	3286	81.1 (79.8–82.5)	3002	87.6 (86.4–88.8)
7	MBS	Parent 2	1999	70.4 (68.4–72.4)	1775	71.6 (69.5–73.7)
7	PBS/RPBS	Parent 2	1999	68.2 (66.1–70.2)	1775	69.0 (66.9–71.2)
7	CLNK	Parent 1	na	na	3002	85.0 (83.8–86.3)
7	CLNK	Parent 2	na	na	1775	59.2 (56.9–61.5)
Study child reporting for themselves						
6	MBS	Study child	na	na	3537	86.6 (85.5–87.7)
6	PBS	Study child	na	na	3537	85.4 (84.2–86.5)
7	CLNK	Study child	na	na	3089	81.2 (79.8–82.5)

ACIR = Australian Childhood Immunisation Register; AEDC = Australian Early Development Census; CI = confidence interval; CLNK = Centrelink income support administrative data; MBS = Medicare Benefits Schedule; na = not applicable; NAPLAN = National Assessment Program – Literacy and Numeracy; PBS = Pharmaceutical Benefits Scheme; RPBS = Repatriation Pharmaceutical Benefits Scheme

Note: Data presented in the table are based on the date of extraction. It is possible to update consent information, over time, through further data cleaning.

59% of Parent 2. Study child consent rates were 87% for MBS, 85% for PBS and 81% for income support administrative data.

Table 5 presents the agreement between the study child's consent to link MBS and PBS data at wave 6 and consent to link income support administrative data. For children with available consent information from both wave 6 for MBS data linkage and wave 7 for income support administrative data linkage (n = 2969), 73.5% consented to both; 15.4% did not consent in wave 6 but did in wave 7; 8.3% consented in wave 6 but not in wave 7; and 2.9% did not consent in both waves.

Children also consented to PBS data linkage at wave 6. Comparing these rates with consent rates for linking income support administrative data at wave 7, we observe that 72.6% consented to both linkages; 15.0% consented to linkage of income support administrative data but not PBS data; 9.1% consented to linkage of PBS data but not income support administrative data; and 3.2% did not consent to both linkages. Seventy-six per cent of children were consistent in their consent decisions in waves 6 and 7 (either consented at both waves 6 and 7 or did not consent at waves 6 and 7).

3.1 Study child's consent to link income support administrative data

At wave 7, 81% of the participating K cohort children provided consent to link income support administrative data (2507 of the participating 3089). At wave 7, 1894 children did not participate in the study as a result of refusal or noncontact.

Each factor was initially examined in logistic regression models. The study child's consent to link income support administrative data was significantly associated with their Aboriginal/Torres Strait Islander status (P = 0.0378), academic problems (P = 0.0003), previous consent to link MBS/PBS data (P < 0.0001), parental smoking status (P = 0.0328), parental argumentativeness (P = 0.0175), parental consent to link income support administrative data (P < 0.0001), equivalised household income (P = 0.0209) and study commitment (P < 0.0001).

In the mutually adjusted logistic regression model (following Peycheva et al. [forthcoming]), the study child's academic problems and parental consent to link income support administrative data were significantly associated with the study child's consent to link income support administrative

Table 5 Agreement between study child's consents at two consecutive waves

Consent for income support administrative data at wave 7	MBS consent at wave 6			PBS consent at wave 6		
	Did not consent, *n* (%)	Did consent, *n* (%)	Total, *N* (%)	Did not consent, *n* (%)	Did consent, *n* (%)	Total, *N* (%)
Did not consent	86 (2.9)	246 (8.3)	332 (11.2)	96 (3.2)	271 (9.1)	367 (12.4)
Did consent	456 (15.4)	2181 (73.5)	2637 (88.8)	446 (15.0)	2156 (72.6)	2602 (87.6)
Total	542 (18.3)	2427 (81.7)	2969 (100.0)	542 (18.3)	2427 (81.7)	2969 (100.0)
Agreement	76.4% (95% CI 74.8–77.9%)			75.9% (95% CI 74.3–77.4%)		

CI = confidence interval; MBS = Medicare Benefits Schedule; PBS = Pharmaceutical Benefits Scheme

data. Study children with academic problems were less likely to consent (OR = 0.71; 95% CI 0.54 to 0.95) than those with no academic problems; and children with parent(s) consenting to link income support administrative data were more likely to consent to data linkage (OR = 10.89; 95% CI 7.64 to 15.53 for single parent consent; and OR = 19.11; 95% CI 12.94 to 28.23 for both parent consent). Overall findings from the model are summarised in Table 6.

3.2 Role of prior data linkage consent

To investigate further the study child's previous consent to link MBS/PBS data (at wave 6) with the study child's consent to link income support administrative data, another model was developed adjusting for all the factors included in Table 6. Children who consented previously to link MBS/PBS data were more likely to consent to link income support administrative data (OR = 1.47; 95% CI 1.09 to 2.00; *P* = 0.0127).

Table 6 Model results for association with study child's consent to link income support administrative data

Factor	Category	Reference category	*P*	Adjusted OR	95% CI
Aboriginal or Torres Strait Islander status	Aboriginal/Torres Strait Islander	Not Aboriginal/Torres Strait Islander	0.7547	1.11	0.59–2.08
Academic problems	Having academic problems	Not having academic problems	**0.0123**	**0.71**	**0.54–0.95**
	Not reported			1.50	0.91–2.47
Parental smoking status	Only one parent currently smokes	Both parents do not smoke	0.6511	1.06	0.81–1.40
	Both parents currently smoke			0.80	0.52–1.24
	Not reported			0.89	0.50–1.57
Parental argumentativeness	One parent reported high argumentativeness	Both parents reported low argumentativeness	0.9852	1.04	0.79–1.36
	Both parents reported high argumentativeness			1.05	0.78–1.42
	Lone-parent families			1.06	0.76–1.49
	Not reported			0.84	0.32–2.21
Parental consent to link income support administrative data	One parent consented	Both parents did not consent	**<0.0001**	**10.89**	**7.64–15.53**
	Both parents consented			**19.11**	**12.94–28.23**
	Parent(s) did not participate			1.10	0.73–1.66
Household income	2nd quintile	1st quintile (lowest quintile)	0.1793	0.82	0.60–1.13
	3rd quintile			1.19	0.84–1.69
	4th quintile			0.84	0.59–1.17
	5th quintile (highest quintile)			0.95	0.66–1.35
Study commitment	Responded to all waves	Marginally attached and responding	0.0711	1.35	0.98–1.87

CI = confidence interval; OR = odds ratio

Note: OR is statistically significant at $Pr(\chi^2 < 0.0001)$ if bolded.

4 Discussion

In this paper, for the first time, we identify factors associated with young people's consent to linkage of income support data in a nationally representative study. To our knowledge, this is the first time consent to data linkage has been studied in young people under the age of 18. We find that young people's consent to data linkage was above 80% in waves 6 and 7, but there was a slight decline in consent rates between the two waves. Moreover, although young people are generally consistent in their consent to data linkage, with 76% either consistently providing consent or not providing consent, 24% changed their consent between waves. In line with prior studies that have shown that data linkage consent is lowest for economic records (Sala et al. 2010, Mostafa 2014, Edwards & Biddle forthcoming), LSAC data linkage consent is also lower for income support data than for medical records.

This paper's central focus was to identify factors influencing the study child's consent to link to income support administrative data using a framework that incorporated novel characteristics (see Figure 1). We considered both of the study child's parents, allowing us to observe whether either parent's characteristics contributed to the study child's consent. We also considered household characteristics, which enabled us to consider the interplay between the study child and their family. Finally, we considered study commitment, which captures response status in prior waves.

We have shown that the study child's consent is significantly associated with multiple characteristics. In the univariate analyses, these included the study child's Aboriginal or Torres Strait Islander status, academic problems, parental smoking status, parental argumentativeness, parental consent to link income support administrative data, household income and study commitment. In multivariate analyses, the study child's academic problems, parental consistency and parental consent to link income support administrative data were statistically significant. We also showed that study child consent is significantly associated with their previous consent to other administrative data linkage.

Parental consent to link income support administrative data was particularly influential in the study child's consent. These findings are consistent with the notion of household contagion, where household members influence each others' decisions, and the findings of Sala et al. (2010) that other members of the household have the greatest influence on a person's propensity to consent to data linkage. Prior consent to data linkage was also a significant factor in current consent to data linkage (even after other factors were taken into account), which is consistent with recent findings that there is a moderate level of consistency in data linkage consent over waves (Jäckle et al. 2018).

Many of the characteristics identified in the univariate analyses were consistent with prior studies that showed that ethnic minorities (e.g. Aboriginal or Torres Strait Islanders; Sala et al. 2010, Baghal et al. 2014, Mostafa 2014, Baghal 2016, Mostafa & Wiggins 2017) and those at lower socioeconomic levels (household income and working hours; Baghal et al. 2014, Mostafa & Wiggins 2017) were associated with lower levels of consent to data linkage. However, consistent with the findings of Peycheva et al. (forthcoming), when a large number of other factors, including demographic characteristics, were considered, many demographic factors were no longer statistically significant. Peycheva et al. (forthcoming) also reported that 24-year-olds' experience with income support was not associated with consent to data linkage. This is consistent with our finding that parental receipt of government benefits was not related to young people's propensity to consent to linkage of income support data.

One of the limitations of the current study is that we did not have measures of trust from young

people or parents, which have been found to be associated with higher levels of consent (Sala et al. 2010, Baghal et al. 2014, Mostafa 2014). However, given that Sala et al. (2010) reported that trust accounts for a small fraction of the variability in consent rates compared with other household members' decisions, this omission is unlikely to have led to a high level of omitted variable bias.

A significant limitation of our study is that it is observational and not experimental. For example, it may be that the household contagion we observed could also reflect other (unobserved) household characteristics. However, it is reassuring that many of our findings accord with experimental studies in terms of differences in rates of consent to data linkage (e.g. Edwards & Biddle [forthcoming]), other observational studies of young people's propensity to consent to data linkage (e.g. Peycheva et al. [forthcoming]), and household panel studies (e.g. Sala et al. [2010]). Another limitation is that we have not specifically modelled the factors associated with data linkage consistency over time. This is an important topic and worthy of further research.

The results have a number of practical implications for survey practitioners. First, it is critical to be mindful of the influence of others in the household when requesting data linkage consent in the context of household surveys. Specific strategies could be employed to increase the influence of others when consent is provided and to offset the influence of consent when consent is denied (within ethical limits). For example, selectively highlighting other family members' consent could be a strategy to increase linkage rates. Second, those who have not responded to all waves may be less likely to consent (although this association was at $P < 0.07$), and may therefore require a greater level of reassurance about the value of the linked information and its security. Third, given that 24% of young people were not consistent in their data linkage consent decision (see Jäckle et al. [2018] for further evidence), for those who decline to consent in a prior wave, it may well be worth asking for consent again (given ethics committee clearance). Finally, on a positive note, although it is well established that consent for data linkage to economic records is lower than for health and education records (e.g. Edwards & Biddle [forthcoming]), the consent rates for this population of young people were still over 80% without substantial targeted encouragement for consent.

Notes

1. A few correlates were consistent across all types of data linkage – specifically, how individuals were treated by government and willingness to answer sensitive questions in a prior wave. Specific to income support data, those who were not working at 17–18 years were less likely to consent (7 years before the request), which conflicts with some other studies that suggest that prior experiences with an institution make a person more likely to consent (e.g. Mostafa & Wiggins 2017). However, prior experiences with an institution were related to consent for data linkage to health records.

References

ABS (Australian Bureau of Statistics) (2008). *Information paper: an introduction to Socio-Economic Indexes for Areas (SEIFA), 2006*, cat. no. 2039.0, ABS, Canberra, www.ausstats.abs.gov.au/Ausstats/subscriber.nsf/0/D729075E079F9FDECA2574170011B088/$File/20390_2006.pdf.

Ahrons CR (1981). The continuing coparental relationship between divorced spouses. *American Journal of Orthopsychiatry* 5:415–428.

AIHW (Australian Institute of Health and Welfare) (2018a). *Medicare Benefits Schedule (MBS) data collection*, AIHW, Canberra, www.aihw.gov.au/about-our-data/our-data-collections/medicare-benefits-schedule-mbs.

——— (2018b). *Pharmaceutical Benefits Scheme (PBS) data collection*, AIHW, Canberra, www.aihw.gov.au/about-our-data/our-data-collections/pharmaceutical-benefits-scheme.

AMHOCN (Australian Mental Health Outcomes and Classification Network) (2005). *Strengths and Difficulties Questionnaire: training manual*, AMHOCN, Parramatta.

Baghal T (2016). Obtaining data linkage consent for children: factors influencing outcomes and potential biases. *International Journal of Social Research Methodology* 19(6):623–643.

——— , Knies G & Burton J (2014). *Linking administrative records to surveys: differences in the correlates to consent decisions*, Understanding Society Working Paper Series 2014-09, Institute for Social and Economic Research, Colchester, United Kingdom.

Bandara D, Gasser C, Jessup K, Renda J, Warren D & Daraganova D (forthcoming). *Factors associated with non-response in Growing Up in Australia: The Longitudinal Study of Australian Children (LSAC)*, Australian Institute of Family Studies, Melbourne.

Bray JR (2001). *Hardship in Australia: an analysis of financial stress indicators in the 1989–99 Australian Bureau of Statistics Household Expenditure Survey*, Occasional Paper 4, Australian Bureau of Statistics, Canberra.

Calderwood L & Lessof C (2009). Enhancing longitudinal surveys by linking to administrative records. In: Lynn P (ed.), *Methodology of longitudinal surveys*, Wiley, Chichester.

CMHR (Centre for Mental Health Research) (2005). *Path Through Life Project PATH 40+*, CMHR, Canberra.

Couper MP (2017). New developments in survey data collection. *Annual Review of Sociology* 43:121–145.

Davis-Kean P, Chambers RL, Davidson LL, Kleinert C, Ren Q & Tang S (2017). *Longitudinal Studies Strategic Review: 2017 report to the Economic and Social Research Council*, UK Economic and Social Research Council.

Edwards B (2014). *Growing Up in Australia*: The Longitudinal Study of Australian Children: entering adolescence and becoming a young adult. *Family Matters* 95:5–14.

——— & Biddle N (forthcoming). Consent to data linkage: experimental evidence on the impact of data linkage requests and understanding and risk perceptions. In: Lynn P (ed.), *Advances in longitudinal survey methodology*, Wiley, London.

Goodman R (2001). Psychometric properties of the strengths and difficulties questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry* 40:1337–1345.

Gray M & Smart D (2009). *Growing Up in Australia*: The Longitudinal Study of Australian Children: a valuable new data source for economists. *Australian Economic Review* 42(3):367–376.

Jäckle A, Beninger K, Burton J & Couper MP (2018). Understanding data linkage consent in longitudinal surveys. In Lynn P (ed.), *Methodology of longitudinal surveys*, Wiley, Chichester.

Kessler RC, Barker PR, Colpe LJ, Epstein JF, Gfoerer JC, Hiripi E, Howes MJ, Normand S-LT, Manderscheid RW, Walters E & Zaslavsky AM (2003). Screening for serious mental illness in the general population. *Archives of General Psychiatry* 60:184–189.

Manfreda L, Bosnjak M, Berzelak J, Haas I & Vehovar V (2008). Web surveys versus other survey modes: a meta-analysis comparing response rates. *International Journal of Market Research* 50(1):79–104.

Mostafa T (2014). *Variation within households in consent to link survey data to administrative records: evidence from the UK Millennium Cohort Study*, CLS Working Paper 2014/8, Centre for Longitudinal Studies, London.

——— & Wiggins R (2017). What influences respondents to behave consistently when asked to consent to health record linkage on repeat occasions? *International Journal of Social Research Methodology* 21:1–16.

Peters TJ (2008). Multifarious terminology: multivariable or multivariate? univariable or univariate? *Paediatric and Perinatal Epidemiology* 22:506.

Pew Research Center (2017). *Americans and cybersecurity*, Pew Research Center, Washington, DC.

Peycheva D, Ploubidis GB & Calderwood L (forthcoming). Determinants of consent to administrative records linkage in longitudinal surveys: evidence from Next Steps. In: Lynn P (ed.), *Methodology of longitudinal surveys*, Wiley, Chichester.

Prensky M (2001). Digital natives, digital immigrants. *On the Horizon* 9(5):1–6.

Sakshaug SW, Couper MP, Ofstedal MB & Weir DR (2012). Linking survey and administrative records: mechanisms of consent. *Sociological Methods Research* 41(4):535–569.

Sala E, Burton J & Knies G (2010). *Correlates of obtaining informed consent to data linkage: respondent, interview, and interviewer characteristics*, ISER Working Paper 2010-28, Institute for Social and Economic Research, Colchester, United Kingdom.

Soloff C, Lawrence D & Johnstone R (2005). *Sample design*, LSAC technical paper 1, Australian Institute of Family Studies, Melbourne.

Statistics Canada (2000). *National Longitudinal Survey of Children and Youth (NLSCY) cycle 3 survey instruments: parent questionnaire*, Statistics Canada, Ottawa.

Wengrzik J, Bosnjak M & Manfreda L (2016). *Web surveys versus other survey modes: a meta-analysis comparing response rates*, paper presented at General Online Research Conference, 2–4 March, Dresden.

CENTRE FOR SOCIAL
RESEARCH & METHODS

+61 2 6125 1279
csrm.comms@anu.edu.au

The Australian National University
Canberra ACT 2601 Australia

www.anu.edu.au