Australian Government
Department of Social Services
Australian Institute of Family Studies

Growing Up in Australia:
The Longitudinal Study of Australian Children (LSAC)
LSAC Technical Paper No. 26

The Longitudinal Study of Australian Children

Wave 9C2 Weighting and non-response

Australian Bureau of Statistics LSAC processing team and the
Australian Bureau of Statistics Household Survey Methodology team

June 2022

Acknowledgements

Growing Up in Australia: The Longitudinal Study of Australian Children is conducted by 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.

This paper is an extension of several previous weighting and non-response technical reports produced by the Australian Institute of Family Studies and the Australian Bureau of Statistics.

Growing Up in Australia: The Longitudinal Study of Australian Children is conducted in partnership between the Australian Government Department of Social Services, the Australian Institute of Family Studies (AIFS) and the Australian Bureau of Statistics (ABS), with advice provided by a consortium of leading researchers from research institutions and universities throughout Australia growingupinaustralia.gov.au.

Technical paper

The Longitudinal Study of Australian Children: LSAC Technical paper No. 26, Wave 9C2 Weighting and Non-Response

Authors: Australian Bureau of Statistics LSAC processing team and the Australian Bureau of Statistics Household Survey Methodology team

For more information, write to:
Longitudinal Studies – Research and Methods
Data Strategy and Development| Data and Evaluation
Australian Government Department of Social Services
PO Box 7576
Canberra Business Centre ACT 2610

Email: LongitudinalStudiesDataAccess@dss.gov.au

ISBN 978-1-76016-255-9 (Online)
ISBN 978-1-76016-256-6 (PDF)

Contents

List of figures

List of tables

Introduction

The Longitudinal Study of Australian Children (LSAC) began in 2004 with a sample of Australian children of two different age cohorts. The study collects data every two years from this sample, subject to attrition from non-response or non-contact.

The sample in the first year was intended to be representative of Australian children in each of the two selected age cohorts, allowing the assessment of developmental outcomes from infancy until middle childhood. Australian children include citizens, permanent residents and applicants for permanent residency (Soloff, Lawrence, & Johnstone, 2005).

The two cohorts of children included in the study were:

- the B cohort, who were aged 0–1 year at the beginning of the study (born between March 2003 and February 2004)
- the K cohort, who were aged 4–5 years at the beginning of the study (born between March 1999 and February 2000).

The first wave of data collection took place in 2004, with subsequent main waves conducted every two years.

Due to COVID, the face-to-face interviews for Wave 9 were cancelled and replaced by two web-based surveys (Wave 9C1 and Wave 9C2). Wave 9C2 of the Longitudinal Study of Australian Children was conducted in 2021 with B cohort children at age 17–18 years and K cohort children at age 21–22 years. Generally, the number of active participants continues to decrease from wave to wave. This is a result of failure to maintain contact, participants opting out of the study (including some instances where the study child/young person is actually deceased), or young people moving out of scope (e.g. moving overseas). Some young people have come back into the sample after missing a wave, if contact can be re-established (e.g. if they return from overseas). There were 18,814 families in the original mailout sample, of which 16,342 were contacted and 10,090 successfully recruited to participate in the study. Of these 10,090 children recruited in the Wave 1 sample, 4,188 young people responded in Wave 9C2, and 3,044 children/young people responded to all waves. In order to maximise the 9C2 response a telephone interview (CATI) was offered, in addition to the web form (CAWI) offered in 9C1, as an alternative approach.

The use of weighting in analysis

Surveys often use probability samples to allow inferences about the population to be drawn. The Longitudinal Study of Australian Children tracks two child cohorts across time, and these were recruited using a probability sample design. Population inference from longitudinal cohorts over time is enabled using two main strategies: retaining a strong proportion of the original selected cohort through effective tracking and follow-up procedures and performing missing data analysis to diagnose and correct for inevitable sample attrition.

The composition of the sample, and thus how well it represents the original population from which the sample was drawn, can be affected by the non-participation of those chosen in the original random selection. The two main mechanisms of non-participation occur during the initial recruitment stage, when persons in the randomly selected sample cannot be contacted or do not agree to participate, and during subsequent waves, through attrition by loss of contact (non-contact), opting out (refusal), or otherwise moving beyond the scope of collection.

This can result in the composition of the active sample being skewed toward or against some demographics, affecting the ability to make inference from the responding sample to the population of interest. If skewed demographics are related to study variables of interest, this can lead to bias when making population inference. Adjusting unit weights to account for attrition can improve the reliability of population inference.

Survey weights are most commonly defined for calculating descriptive statistics and are essential in making accurate inferences from sample frequencies, particularly when missing data are not missing at random (Little & Rubin, 1987). Examples of descriptive statistics in a longitudinal study include the proportion of the children achieving a certain level of educational success or the proportion of the cohort improving on their educational success in the time span between waves.

Longitudinal analytic statistics; for example, the strength of correlations of modelled predictors for children improving on their educational success over time, can also be biased if missing participants behave differently to those remaining in the study. Some longitudinal analysis methods reduce bias by applying survey weights,

while other methods reduce bias by including variables related to response propensity in the modelling process (Pfeffermann, 1993). Here, we highlight that the responsibility lies with the analyst to ensure that their methods are robust against the possible presence of bias due to missing data (Fairclough, 2010).

With this in mind, this paper describes the process of calculating weights for Wave 9C2 of the Longitudinal Study of Australian Children, with a focus on the treatment of bias. We encourage data users to consider the use of survey weights in their analysis if reconfiguring the data to reflect the population is important for the analysis they are undertaking. We also offer a timely reminder to users that LSAC is based on a clustered sample design using a primary sampling unit of postcode, and that this variable should be used when conducting statistical tests to avoid overstating significance.

Summary of sample design properties

Full details about the LSAC sample design can be found in Soloff and colleagues (2005).

Table 1 provides a summary for reference.

Table 1: LSAC sample design properties

Property	Description
Scope (the population about which inference is to be made)	Two cohorts of children (the B cohort who were 0–1 year and the K cohort who were 4–5 years old during 2004, the Wave 1 recruitment year). The scope excluded very remote areas of Australia.
Coverage (the population represented by the active participating sample)	For Wave 1 recruitment: The subset of Wave 1 scope who had contact records available through Medicare, who could be contacted and who agreed to participate in LSAC. For subsequent waves: The subset of Wave 1 coverage who could be contacted. This included tracking address changes and re-recruitment after missing waves where possible, including cases of temporarily moving overseas.
Stratification (division of population into cells from which sample was drawn)	Cells of state x capital city/balance of state x large/small postcode
Selection frame (from which children were selected and contact details obtained)	List frame of Medicare records for children in scope
Sample design	Multi-stage cluster sampling
Selection unit(s)	Stage 1 Unit: Postcode Stage 2 Unit: One cluster of dwellings within postcode Stage 3 Unit: Children in dwellings in cluster
Reporting unit(s)	Parent 1, Parent 2, Child (when old enough), Interviewer, Child care worker, Teacher, Parent Living Elsewhere
Tabulation unit	Child
Selected sample size and fraction	Approximately 10,000 per cohort; approximately 4% of each cohort population
Recruited sample size and fraction at Wave 1	Approximately 5,000 per cohort; approximately 2% of each cohort population
Design effects (factors by which variance is higher under cluster sampling as compared to simple random sampling)	Approximately 90% of LSAC variables have a design effect below 1.5 as stated in Wave 1 Weighting Paper.

Responding units

Table 2 reflects information, also provided in the Data User Guide, about what has been included as a responding record for each wave and cohort and therefore in scope for weighting.

In the earlier waves, Parent 1 was the main respondent, whereas over waves the Study Child (referred to as the Young Person in later waves) started to become a key respondent on their own. Therefore, records were considered responding if we were able to conduct an interview with either Parent 1 or the Study Child. In

Waves 1–7 confirmation of a Parent 2 or a Parent Living Elsewhere was contingent on having had an interview with Parent 1. However, in Wave 8, parent records were generated if the Study Child did not object to a particular parent being interviewed, regardless of whether the Study Child was interviewed or not. In Waves 8 and 9C1, records were considered responding if we were able to conduct an interview with the young person or any of the parental figures. In Wave 9C2, records were considered responding if the Young Person or Parent 1 completed a CAWI or CATI. The Parent 2 and Parent Living Elsewhere were not approached for an interview in Wave 9C2.

Table 2: Responding units by wave

	1	2	3	4	5	6	7	8	9C1	9C2
B	P1	P1	P1	P1	P1 or SC	P1 or SC	P1 or SC	P1 or SC	SC or P1 orP2 or PLE	SC or P1
K	P1	P1	P1	P1 or SC	P1 or SC	P1 or SC	P1 or SC	SC or P1 or P2 or PLE	SC or P1 or P2 or PLE	SC or P1

Summary of weighting in Waves 1–9C1

Weights for Wave 1 were calculated beginning with the inverse probability of selection for each child and then adjusting these weights to align to known population benchmarks (Soloff, Lawrence, Mission, & Johnstone, 2006). A complex variant on the method of post-stratification was used whereby alignment was achieved for row-and-column totals of key benchmark demographics but not all cross-classified cells. This method has variously been termed incomplete post-stratification or calibration to marginal benchmarks and is useful when complete post-stratification would subdivide the sample too finely and lead to model overfitting and large weight changes (Akaike, 1974). Benchmarks for children in the B and K cohorts for each state by capital city/rest of state area were drawn from the ABS Estimated Resident Population as at March 2004, and benchmarks for households by language spoken at home and mother's education level within each region were generated using proportions taken from the 2001 Census.

Weights for Waves 2–9C1 were calculated by adjusting previous wave weights for differential sample attrition in two stages (Cusack & Defina, 2014; Sipthorp & Daraganova, 2011; Sipthorp & Misson, 2007, 2009; Usback, 2018). At the first stage, a modelled response propensity factor was applied; at the second, the weights were adjusted to preserve stratum totals. Extreme weights were capped as a form of outlier treatment to avoid any particular child contributing much more than other children in the sample to a weighted estimate, because this can potentially lead to volatile statistics if any such child has unusual characteristics. In Wave 9C1, sample weight bounds were revised from [0.33, 3.5] to [0.28, 4.5] in light of low response rates (due to COVID-19). Expanding the weights cap allows the proportion of capped units in previous waves to be consistent.

In each wave, a population weight is calculated that adds up to the number of children in the Wave 1 population and in the corresponding age group and year, and a sample weight is calculated that adds up to the number of children in the sample for that wave. The population weight conceptually represents the number of children in the population represented by each child in the sample when creating weighted estimates. The sample weight can be used as a measure of the representativeness of each child compared to the others in the sample. The sample weights are equal to the population weights multiplied by the sampling fraction.

In Waves 2–4, weights were produced for every combination of response to individual waves. In Wave 5, this was simplified to a concise set of eight weights: each cohort has a longitudinal weight (both sample and population weights), and a cross-sectional weight (both sample and population weights). The longitudinal and cross-sectional weights are produced for different combinations of response:

- The **longitudinal weights** are defined for the sample responding to all waves up to and including the current wave, and involve an adjustment made for each new wave response. Longitudinal weights are most suitable for analysis that makes use of data from many time periods.
- The **cross-sectional weights** are defined for the sample responding only to the most recent wave, irrespective of response to all or some of the intervening waves since Wave 1. Cross-sectional weights are most suitable for analysis that makes use only of the current data.

For more information on weighting methods from previous waves, please refer to the technical papers available on the *Growing Up in Australia* website.

Wave 9C2 weighting method

This section contains a brief description of the method used to create weights for Wave 9C2 data. The method is largely unchanged from Wave 9C1.

The weighting process for LSAC is in two stages. First, the response propensity modelling adjustment is applied to correct for attrition between waves. Second, the stratum adjustment is applied to realign weight totals with known totals from the original sample. Both stages contribute to non-response bias reduction.

Longitudinal weights are calculated by taking the longitudinal weight from the previous wave of the study and adjusting for any additional non-response in the current wave.

For calculation of cross-sectional weights, the final weights used in Wave 1 are adjusted for all additional non-responses in the current wave – regardless of whether a unit responded to Waves 2–9C1.

Initial weights

The final weights of a previous wave are carried forward to become the initial weights for the next wave.

- For Wave 9C2 longitudinal weights (which apply to those who have responded to all Waves 1, 2, 3, 4, 5, 6, 7, 8, 9C1 and 9C2), the initial weight for children in Wave 9C2 is the final longitudinal corrected weight from Wave 9C1.
- For Wave 9C2 cross-sectional weights (which apply to all of those who responded to Wave 9C2), the initial weight for children in Wave 9C2 is the final weight from Wave 1.

Response propensity modelling

The purpose of this step is to adjust for differential non-response by particular demographic groups or individual characteristics that may have higher or lower sample attrition than average. This is done by modelling the response propensity using logistic regression (Little, 1986), using the dataset of respondents and non-respondents together, and using past wave survey responses as regressors. The modelled propensity is then used as a weight adjustment factor. For example, if a unit's response propensity is modelled at 90% then its response propensity adjusted weight is calculated as its initial weight divided by 0.9.

Selection of covariates for logistic regression non-response adjustment

The method for selection of covariates to use in the response propensity model is largely unchanged from Wave 9C1. A stepwise model selection process is used that considers all possible covariates for the response propensity model (list of covariates provided in Appendix E).

This stepwise process calculates the score of chi-square statistics of covariates not in the model and adds the largest covariate not yet in the model. If any covariates are no longer found to be significant ($p < 0.05$) then they are removed from the model. This model selection process resulted in a shortlist of variables to consider adding to the Wave 9C2 model.

The response propensity model for Wave 9C2 was then re-run on the shortlisted variables together with the variables used for the derivation of Wave 9C1 weights. Variables that were no longer useful or significant ($p > 0.1$) were removed from the model. Variables used in the Wave 9C1 derivation that were still useful predictors in the Wave 9C2 model were maintained where possible to achieve consistency over time.

Wave 1 variables used in the B cohort cross-sectional weight model

- Parent 1: age
- Parent 2: age
- Mother: school completion
- Mother: main language spoken at home (other than English)
- Parent 1: self-complete data present
- Parent 2: self-complete data present
- Parent 1: rents home (renting home indicator)
- Parent 1: Australia as country of birth
- SEIFA Education and Occupation
- Number of siblings of study child in household
- Study child: sex

Wave 1 variables used in the K cohort cross-sectional weight model

- Parent 1: age
- Mother: school completion
- Parent 1: self-complete data present
- Parent 2: self-complete data present
- Parent 1: rents home (renting home indicator)
- Parent 1: English as main language at home
- SEIFA Education and Occupation
- Stratum
- Study child: sex

Wave 9C1 variables used in the B cohort longitudinal weight model

- Parent 1: age
- Parent 1: completed CAWI (new)
- Parent 2: completed CAWI
- Study child: current level of achievement in studies (new)

Wave 9C1 variables used in the K cohort longitudinal weight model

- Study child: number of people living with study child
- Study child: sex
- Parent 1: employment status (new)

Model significance tests of the data items used in the above models can be found in Appendix C.

Odds ratio estimates for the data items used in the above models can be found in Appendix D.

A list of the variables considered in the selection of covariates for the response propensity models can be found in Appendix E.

Stratum weight adjustment

The purpose of this step is to use weighting to re-align the sample composition within each stratum as at Wave 1, and to re-align the sum of sample weights to be equal to the number of original participants in the first wave. The original selections were done by dividing each state into a capital city statistical division versus rest of state and then into groups of large or small postcodes. These are the original strata.

This adjustment accounts for some non-responses not already adjusted in the model and ensures consistent estimates at the stratum level over time.

This stratum weight adjustment is also known as post-stratification or calibration to benchmarks. There is a separate adjustment factor calculated for each stratum based on the sum of the response propensity adjusted weights compared to the benchmark of the count of children within that stratum, subject to individual sample weights not exceeding the lower weight cap of 0.28 or the upper weight cap of 4.5. This process of calculating the weight adjustment for each unit to satisfy the benchmark specified while simultaneously satisfying the weight caps specified is achieved iteratively through the ABS SAS implementation of the generalised regression estimator (GREGWT). To avoid larger adjustments of weight in strata with a small number of responding children, several strata were collapsed with other strata within the same state for the stratum weight adjustment.

Weight capping

Weight capping is the process of limiting extreme values of weights for records that would otherwise have a large influence on estimates and calculations. Extreme weights can result from the logistic regression response propensity modelling step if a respondent's predicted chance of responding is very low, leading to a large weight adjustment. Weight capping is a robust form of automatic treatment of extreme values for weights, improving the variance characteristics of any analysis performed at the expense of a slight reduction in contribution for some respondent groups (i.e. a slight risk of bias).

The weight caps are applied during the stratum weight adjustment step to ensure that any large response propensity adjusted weights are adjusted back to a reasonable level.

The lower cap of 0.28 and the upper cap of 4.5 remain the same for Wave 9C2. The sample weight bounds were updated in Wave 9C1 because of the low response rates (mainly due to COVID-19).

More detail on the number of units now appearing at the caps can be seen in Tables 6 and 7 in the next section of this paper.

Further characteristics of response across waves

Reacquisition of sample from previous waves

In this context, the reacquisition of sample refers to gaining a full response from a participant who was not considered responding in a previous wave. For the B cohort, out of 3,090 that did not respond to Wave 9C1, 847 responded to Wave 9C2. Out of 3,370 the that did not respond to at least one of Waves 2, 3, 4, 5, 6, 7, 8 or 9C1, 1,076 responded to Wave 9C2.

For the K cohort, out of 3,194 that did not respond to Wave 9C1, 816 responded to Wave 9C2. Out of the 3,461 that did not respond to at least one of Waves 2, 3, 4, 5, 6, 7, 8 or 9C1, 1,031 responded to Wave 9C2.

Table 3 shows those who have responded after previously being a 'non-responder' in a previous wave (sample reacquisition).

Table 3: Sample reacquisition for Waves 3, 4, 5, 6, 7, 8, 9C1 and 9C2

Cohort	Resp. Wave 3, not Wave 2	Resp. Wave 4, not Wave 3	Resp. Wave 5, not Wave 4	Resp. Wave 6, not Wave 5	Resp. Wave 7, not Wave 6	Resp. Wave 8, not Wave 7	Resp. Wave 9C1, not Wave 8	Resp. Wave 9C2, not Wave 9C1
B	133	135	129	89	124	134	118	847
K	135	119	94	77	120	301	70	816

For the B cohort, there were 847 units that responded to Wave 9C2 that did not respond to Wave 9C1. Of these 847 units, there were 624 units where the study child was interviewed. For the K cohort, of the 816 units that responded to Wave 9C2 that did not respond to Wave 9C1, there were 516 units where the study child was interviewed. Refer to Table 2 for a definition of what is considered a responding unit.

Total responding sample for each wave

The fully responding sample at each wave drives the calibration and, hence, the weighting process.

Observe Tables 4 and 5 for updated counts.

Table 4: Sample counts for the B cohort

Wave	1	2	3	4	5	6	7	8	9C1	9C2
Cross-sectional response	5,107	4,606	4,386	4,242	4,085	3,764	3,381	3,127	2,017	2,688
Longitudinal response	-	4,606	4,253	3,997	3,758	3,441	3,028	2,722	1,737	1,612
Cross-sectional attrition rate (%)	-	9.8	14.1	16.9	20.0	26.3	33.8	38.8	60.5	47.4
Longitudinal attrition rate (%)	-	9.8	7.7	6.0	6.0	8.4	12.0	10.1	36.2	7.2

Table 5: Sample counts for the K cohort

Wave	1	2	3	4	5	6	7	8	9C1	9C2
Cross-sectional response	4,983	4,464	4,331	4,169	3,956	3,537	3,089	3,037	1,789	2,463
Longitudinal response	-	4,464	4,196	3,940	3,682	3,276	2,792	2,510	1,522	1,432
Cross-sectional attrition rate (%)	-	10.4	13.1	16.3	20.6	29.0	38.0	39.1	64.1	50.6
Longitudinal attrition rate (%)	-	10.4	6.0	6.1	6.5	11.0	14.8	10.1	39.4	5.9

- Cross-sectional response – number of children who responded to that particular wave.
- Longitudinal response – number of children who have responded to all waves up to and including that particular wave; that is, fully responding to each wave since Wave 1.
- Cross-sectional attrition rate (%) – those not responding to that particular wave as a percentage of the Wave 1 cross-sectional response.
- Longitudinal attrition rate (%) – those not responding to the current wave but having responded to all waves beforehand, as a percentage of the previous wave's longitudinal response.

Number of children with weight at cap

Tables 6 and 7 show the number of children with a sample weight at the lower cap of 0.28 and upper cap of 4.5 by cohort and by type of weight.

For the B cohort, the number of units at the upper cap has decreased from 34 in Wave 9C1 to 24 for the cross-sectional weight; and decreased from 27 in Wave 9C1 to 22 for the longitudinal weight.

Table 6: Counts of capped sample weights for Wave 9C2 – B cohort

	Cross-sectional		Longitudinal	
State	Lower cap (0.28)	Upper cap (4.5)	Lower cap (0.28)	Upper cap (4.5)
NSW	0	10	0	8
Vic.	0	6	0	4
Qld	4	3	0	5
SA	0	3	1	2
WA	0	1	1	2
Tas.	0	1	0	1
NT	11	0	10	0
ACT	0	0	1	0
Australia	**15**	**24**	**13**	**22**

For the K cohort, the number of units at the upper cap increased from 11 in Wave 9C1 to 15 for the cross-sectional weight; and decreased from 10 in Wave 9C1 to 5 for the longitudinal weight.

Table 7: Counts of capped sample weights for Wave 9C2 – K cohort

	Cross-sectional		Longitudinal	
State	Lower cap (0.28)	Upper cap (4.5)	Lower cap (0.28)	Upper cap (4.5)
NSW	0	5	0	0
Vic.	0	6	0	1
Qld	1	4	5	3
SA	0	0	0	0
WA	0	0	2	1
Tas.	9	0	4	0
NT	15	0	6	0
ACT	1	0	2	0
Australia	**24**	**15**	**19**	**5**

Conclusion

Some sample was regained this wave; with the responding sample around 2,500 for the K cohort and around 2,700 for the B cohort. The longitudinal dataset presents a rich source of information about Australian children. The response propensity models identify which characteristics of the sample were related to their probability of response. The weights developed help to correct for different response patterns, allowing users to better analyse the data and draw more accurate conclusions about the population, being the two cohorts of children (the B cohort who were 0–1 year and the K cohort who were 4–5 years old during 2004, the Wave 1 recruitment year).

The weight capping ensures that no unit contributes too much or too little to any analysis using these data.

Bibliography

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*(6), 716–723.

Australian Bureau of Statistics. (2013). *Australian Demographic Statistics*, Sep. 2012. Canberra: Australian Bureau of Statistics.

Australian Institute of Family Studies (Ed.). (2013). *The Longitudinal Study of Australian Children annual statistical report 2012*. Melbourne: Australian Institute of Family Studies.

Bell, P. (2000). *Weighting and standard error estimation for ABS household surveys*. Australian Bureau of Statistics Methodology Advisory Committee Paper. Canberra: Australian Bureau of Statistics.

Cusack, B., & Defina, R. (2014). *LSAC Technical Paper No. 10: Wave 5 weighting and non-response*. Melbourne: Australian Institute of Family Studies.

Engle, R. (1983). Wald, Likelihood Ratio, and Lagrange multiplier tests in Econometrics. In Z. Griliches & M. D. Intriligator (Eds.), *Handbook of Econometrics II* (pp. 796–801). Elsevier.

Fairclough, D. L. (2010). *Design and analysis of quality of life studies in clinical trials*. Boca Raton, FL: Chapman and Hall/CRC.

Holt, D., & Smith, T. M. F. (1979). Post-stratification. *Journal of the Royal Statistical Society Series A, 142*, 33–46.

Little, R. J. A. (1986). Survey nonresponse adjustments for estimates of means. *International Statistical Review, 54*, 139–157.

Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data*. New York: Wiley.

Pfeffermann, D. (1993). The role of sampling weights when modelling survey data. *International Statistical Review, 61,* 317–337.

Sarndal, C. E., Swensson, B., & Wretman, J. H (1992). *Model assisted survey sampling*. New York: Springer-Verlag,

Sipthorp, M., & Misson, S. (2007). *LSAC Technical Paper No. 5: Wave 2 weighting and non-response*. Melbourne: Australian Institute of Family Studies.

Sipthorp, M., & Misson, S. (2009). *LSAC Technical Paper No. 6: Wave 3 weighting and non-response.* Melbourne: Australian Institute of Family Studies.

Sipthorp, S., & Daraganova, G. (2011). *LSAC Technical Paper No. 9: Wave 4 weights*. Melbourne: Australian Institute of Family Studies.

Soloff, C., Lawrence, D., & Johnstone, R. (2005). *LSAC Technical Paper No. 1: Sample design.* Melbourne: Australian Institute of Family Studies.

Soloff, C., Lawrence, D., Misson, S., & Johnstone, R. (2006). *LSAC Technical Paper No. 3: Wave 1 weighting and non-response*. Melbourne: Australian Institute of Family Studies.

Swets, J. A. (1973). The relative operating characteristic in psychology. *Science, 182*, 990–1000.

Usback, S. (2018). *LSAC Technical Paper No. 20: Wave 7 weighting and non-response*. Melbourne: Australian Institute of Family Studies

Appendix A: Glossary of terms and abbreviations

This paper uses many technical terms, some of which are not consistently used across the fields of longitudinal studies and sample designs. We offer a brief glossary as a guide to how the terms are used in this paper.

Term	Definition
ABS	Australian Bureau of Statistics
Attrition	Process of sample size shrinking over time due to any mechanism
Cohort	Sample with a particular characteristic, e.g. B cohort ages 0–1 year in first wave
Coverage	Population represented by the remaining active participants
Cross-sectional	Pertaining to a statistic at one time point, typically broken down by characteristics at that time point
Design effect	Penalty factor to variance due to sample tending to be similar within selected postcode clusters
Estimation	Process of calculating a descriptive statistic from sample using weight, acknowledging the presence of sampling error
F2F	Face-to-face
Longitudinal	Pertaining to a statistic involving many time points, typically with a focus on evolution of participants over time
LSAC	Longitudinal Study of Australian Children
Missing data	Data absent either from non-response or partial response
Non-response	Failure to acquire survey response due to non-contact or refusal (opt-out)
P1	Parent 1, generally the child's mother
P2	Parent 2, the child's second parent
PLE	Parent Living Elsewhere
Partial response	Acquisition of data for some study modules but not others
Post-stratification	Process of dividing population into post-strata for weighting
Recruited sample	Subset of selected sample who agreed to participate in Wave 1
Response propensity	Chance that a particular individual or group will respond to a given wave
Respondent	or Participant or Active Participant: Any child (family) active in the study
Selected sample	Selection of children (families) approached at time of Wave 1 recruitment
Stratification	Process of dividing population into strata for selection
Stratum (Strata)	Cell(s) of population from which set number of children selected in sample
Study variable	Any variable collected in the study that data users wish to analyse
Weight	Value for a respondent to correct, up or down, for representativeness based on characteristics of responding sample

Appendix B: Description of Wave 9C2 weights

Table B1: Description of Wave 9C2 weights

SAS name	Cohort	Type	Waves cases responded to
i2weight	B	Population	1 & 9C2
i2weights	B	Sample	1 & 9C2
bcdefghi1i2wt	B	Population	1, 2, 3, 4, 5, 6, 7, 8, 9C1 & 9C2
bcdefghi1i2wts	B	Sample	1, 2, 3, 4, 5, 6, 7, 8, 9C1 & 9C2
k2weight	K	Population	1 & 9C2
k2weights	K	Sample	1 & 9C2
defghijk1k2wts	K	Population	1, 2, 3, 4, 5, 6,7, 8, 9C1 & 9C2
defghijk1k2wt	K	Sample	1, 2, 3, 4, 5, 6, 7, 8, 9C1 & 9C2

Appendix C: Logistic regression models: type 3 analysis of effects

Note that where a response was not obtained to a variable, this was included in the model.

Table C1: B cohort – cross-sectional weights

Variable name	Description	DF[a]	Wald Chi-Square[b]	Pr > ChiSq
acnfseo	SEIFA Index of Education & Occupation	1	15.8	<.0001
af03m2	Parent 1 age	1	21.2	<.0001
af03m3	Parent 2 age	1	8.2	0.0042
ansib	Number of siblings of study child in household	1	25.9	<.0001
af11am	Mother's language other than English spoken at home	1	21.5	<.0001
afd08m1	Mother's highest level of schooling completed	5	104.4	<.0001
aho04a3b	Parent 1 rents home	2	20.8	<.0001
ap1scd	Parent 1 self-completed questionnaire returned	1	23.5	<.0001
ap2scd	Parent 2 self-completed questionnaire returned	2	11.0	0.0042
zf09m2	Parent 1 born in Australia	1	6.0	0.0146
zf02m1	Study child: sex	1	6.4	0.0115

Notes: [a] Degrees of Freedom; [b] Wald Chi-Square is computed by squaring the ratio of the parameter estimate divided by its standard error estimate.

Table C2: B cohort – longitudinal weights

Variable name	Description	DF	Wald Chi-Square	Pr > ChiSq
hf03hp1	Parent 1 age (Wave 8)	1	8.5	0.0036
i1p1cawi	Parent 1 completed CAWI	1	22.9	<.0001
i1pc82c7c	Study child: current level of achievement in studies	6	51.1	<.0001
i1p2cawi	Parent 2 completed CAWI	1	2.9	0.09

Table C3: K cohort – cross-sectional weights

Variable name	Description	DF	Wald Chi-Square	Pr > ChiSq
cf03m2	Parent 1 age	1	44.8	<.0001
cp1scd	Parent 1 self-completed questionnaire returned	1	20.2	<.0001
cp2scd	Parent 2 self-completed questionnaire returned	2	62.9	<.0001
cho04a3b	Parent 1 rents home	2	15.4	0.0005
cfd08m1	Mother's highest level of schooling completed	7	101.6	<.0001
cf11m2	Parent 1 English as main language at home	1	12.5	0.0004
ccnfseo	SEIFA Index of Education & Occupation	1	32.3	<.0001
stratum	Stratum	21	38.8	0.0105
zf02m1	Study child: sex	1	17.2	<.0001

Table C4: K cohort – longitudinal weights

Variable name	Description	DF	Wald Chi-Square	Pr > ChiSq
k1fd36c	Number of people living with study child	1	25.8	<.0001
k1pw30a1a	Parent 1 employment status	7	36.6	<.0001
zf02m1	Study child: sex	1	8.1	0.0044

Appendix D: Odds ratio estimates for variables in Wave 9C2 response propensity models

These odds ratios show different categories of variables included in the model.

Variable categories can be collapsed or re-parameterised as part of the weighting process. Small categories may have been collapsed during this process. Re-parameterisation is the re-labelling of modal categories. The odds ratios are calculated using the maximum category. Re-labelling the reference category to be the maximum makes the results easier to interpret. All the information needed about changes in categories is contained in the description column of the tables below. Please do not compare these results against the Data Dictionary. These changes also apply to the tables in Appendix H.

Table D1: Odds ratio estimates for B cohort – cross-sectional weight

Effect	Description	Point estimate	95% Wald confidence interval	
acnfseo	SEIFA Index of Education & Occupation	1.002	1.001	1.002
ansib	Number of siblings of study child in household	0.853	0.803	0.907
af03m2	Parent 1 age	1.036	1.02	1.051
af03m3	Parent 2 age	1.02	1.006	1.034
af11am 0 vs 1201	Mother speaks English at home vs other language	0.626	0.514	0.763
afd08m1 6	**Mother's highest level of schooling completed Year 12 or equivalent – reference category**			
afd08m1 -2 vs 6	Mother's school completion not asked, refused or don't know	2.476	0.461	13.286
afd08m1 2 vs 6	Mother completed Year 11 or equivalent	3.321	1.867	5.909
afd08m1 3vs 6	Mother completed Year 10 or equivalent	1.972	1.086	3.583
afd08m1 4 vs 6	Mother completed Year 9 or equivalent	1.743	0.968	3.139
afd08m1 5 vs 6	Mother completed Year 8 or below	1.064	0.532	2.126
aho04a3b 2	**Parent 1 does not rent home – reference category**			
aho04a3b -4 vs 2	Refusal or don't know if Parent 1 rents home	0.549	0.09	3.35
aho04a3b 1 vs 2	Parent 1 rents home	0.724	0.63	0.832
ap1scd 0 vs 1	Parent 1 did not return self-completed questionnaire vs Parent 1 did return self-completed questionnaire	0.55	0.432	0.7
ap2scd 1	**Parent 2 self-complete questionnaire returned – reference category**			
ap2scd -9 vs 1	No Parent 2 in household	0.847	0.52	1.379
ap2scd 0 vs 1	Parent 2 did not return self-completed questionnaire	0.695	0.56	0.863
zf02m1 1 vs 2	Study child sex: male vs female	0.858	0.763	0.966
zf09m2 0 vs 1101	Parent 1: born in Australia vs born elsewhere	0.811	0.686	0.96

Table D2: Odds ratio estimates for B cohort – longitudinal weight

Effect	Description	Point estimate	95% Wald confidence interval	
hf03hp1	Parent 1 age	1.055	1.018	1.093
i1p1cawi -9 vs 1	Parent 1 did not respond vs Parent 1 completed CAWI	0.394	0.269	0.577
i1p2cawi -9 vs 1	Parent 2 did not respond vs Parent 2 completed CAWI	0.7	0.463	1.057
i1pc82c7c 5	**Study child current level of achievement in studies very low – reference category**			
i1pc82c7c -9 vs 5	Study child not in education	0.722	0.291	1.79
i1pc82c7c -5 vs 5	Study child skipped question	1.499	0.158	14.199
i1pc82c7c 1 vs 5	Study child very high achievement	15.84	1.838	136.549
i1pc82c7c 2 vs 5	Study child high achievement	4.555	1.554	13.356
i1pc82c7c 3 vs 5	Study child average achievement	2.015	0.778	5.216
i1pc82c7c 4 vs 5	Study child low achievement	3.389	0.968	11.868

Table D3: Odds ratio estimates for K cohort – cross-sectional weight

Effect	Description	Point estimate	95% Wald confidence interval	
ccnfseo	SEIFA Index of Education & Occupation	1.003	1.002	1.004
cf03m2	Parent 1 age	1.04	1.028	1.053
cf11m2 0 vs 1201	Parent 1 speaks English at home vs or other language	0.725	0.606	0.866
cp1scd 0 vs 1	Parent 1 did not return self-completed questionnaire vs Parent 1 did return self-completed questionnaire	0.578	0.455	0.734
cp2scd 1	**Parent 2 self-complete questionnaire returned – reference category**			
cp2scd -9 vs 1	No Parent 2 in household	0.508	0.415	0.622
cp2scd 0 vs 1	Parent 2 did not return self-completed questionnaire	0.52	0.418	0.645
cho04a3b 2	**Parent 1 does not rent home – reference category**			
cho04a3b -2 vs 2	Refusal or don't know if Parent 1 rents home	0.649	0.1	4.208
cho04a3b 1 vs 2	Parent 1 rents home	0.745	0.643	0.864
cfd08m1 7	**Mother's highest level of schooling completed Year 12 or equivalent – reference category**			
cfd08m1 -9 vs 7	Mother's school completion not asked	0.458	0.212	0.987
cfd08m1 -4 vs 7	Mother's school completion refused or don't know	0.421	0.06	2.938
cfd08m1 2 vs 7	Mother completed Year 11 or equivalent	0.669	0.556	0.805
cfd08m1 3 vs 7	Mother completed Year 10 or equivalent	0.529	0.452	0.619
cfd08m1 4 vs 7	Mother completed Year 9 or equivalent	0.345	0.245	0.486
cfd08m1 5 vs 7	Mother completed Year 8 or below	0.42	0.271	0.651
cfd08m1 6 vs 7	Mother never attended school	0.674	0.138	3.297
stratum 81	**ACT – reference category**			
stratum 11 vs 81	NSW Met	0.573	0.366	0.897
stratum 13 vs 81	NSW Xmet large	0.888	0.552	1.429
stratum 14 vs 81	NSW Xmet small	0.658	0.326	1.33
stratum 21 vs 81	VIC Met large	0.618	0.393	0.974
stratum 22 vs 81	VIC Met small	0.38	0.131	1.108
stratum 23 vs 81	VIC Xmet large	0.678	0.406	1.132
stratum 24 vs 81	VIC Xmet small	0.621	0.333	1.155

Table continued over page ⟶

Effect	Description	Point estimate	95% Wald confidence interval	
stratum 31 vs 81	QLD Met	0.555	0.342	0.901
stratum 33 vs 81	QLD Xmet large	0.481	0.297	0.779
stratum 34 vs 81	QLD Xmet small	0.656	0.339	1.27
stratum 41 vs 81	SA Met large	0.677	0.408	1.126
stratum 43 vs 81	SA Xmet large	0.555	0.28	1.099
stratum 44 vs 81	SA Xmet small	0.734	0.256	2.1
stratum 51 vs 81	WA Met large	0.623	0.382	1.015
stratum 52 vs 81	WA Met small	0.726	0.229	2.301
stratum 53 vs 81	WA Xmet large	0.987	0.531	1.836
stratum 54 vs 81	WA Xmet small	0.612	0.292	1.28
stratum 61 vs 81	TAS Met	0.897	0.436	1.846
stratum 63 vs 81	TAS Xmet	0.788	0.416	1.496
stratum 71 vs 81	NT Met	0.825	0.379	1.795
stratum 73 vs 81	NT Xmet small and NT Xmet large	0.802	0.36	1.783
zf02m1 1 vs 2	Study child sex: male vs female	0.776	0.688	0.875

Table D4: Odds ratio estimates for K cohort – longitudinal weight

Effect	Description	Point estimate	95% Wald confidence interval	
k1fd36c	Number of people living with study child	1.483	1.274	1.726
k1pw30a1a 6	**Parent 1 Not employed, not seeking employment – reference category**			
k1pw30a1a -9 vs 6	Parent 1 employment status not asked	0.428	0.176	1.039
k1pw30a1a -5 vs 6	Parent 1 employment status not answered	0.803	0.085	7.571
k1pw30a1a 1 vs 6	Parent 1 full-time employee	1.847	0.668	5.11
k1pw30a1a 2 vs 6	Parent 1 part-time employee	3.302	0.975	11.183
k1pw30a1a 3 vs 6	Parent 1 self-employed	1.445	0.389	5.373
k1pw30a1a 4 vs 6	Parent 1 unpaid family worker	0.599	0.063	5.694
k1pw30a1a 5 vs 6	Parent 1 unemployed – seeking employment	1.929	0.218	17.05
zf02m1 1 vs 2	Study child sex: male vs female	0.52	0.332	0.816

Appendix E: Data items considered for response propensity models

Table E1: Wave 1 data items considered for B cohort – cross-sectional weight

Variable name	Variable label
acnfsad	0/1—Home—SEIFA Advantage/Disadvantage
acnfseo	0/1—Home—SEIFA Education & Occupation
acnfser	0/1—Home—SEIFA Economic Resources
af01am	0/1—M@0/1—Present for wave
af01m3	0/1—P2@W1—Present for wave
af03m2	0/1—P1@W1—F2F A4—Age
af03m3	0/1—P2@W1—F2F A4—Age
af11am	0/1—M@0/1—F2F A12—Main language spoken at home
af11m1	0/1—SC—F2F A12—Main language spoken at home
af11m2	0/1—P1@W1—F2F A12—Main language spoken at home
afd08a1	0/1—P1—F2F H3—School completion
afd08m1	0/1—M—F2F H3—School completion
afd11m2	0/1—M—F2F H10—Proficiency in spoken English
aho04a3b	0/1—P1—F2F L4—Rent home
aho04a5	0/1—P1—F2F L5—Housing tenure
aho09a1a1	0/1—P1—F2F L11—Safe neighbourhood
anpeople	0/1—No. people in household
ansib	0/1—No. siblings of SC in household
ap1scd	0/1—Parent 1 self-completed data present
ap2	0/1—SC has 2 parents in the home
ap2scd	0/1—Parent 2 self-completed data present
zf02m2	P1@W1—F2F A3—Sex
zf09m2	P1@W1—F2F A10—Country of birth
zf12m1	SC—F2F A13—Indigenous status
zf12m2	P1@W1—F2F A13—Indigenous status
zf02m1	SC - F2F A3 - Sex

Table E2: Wave 9C1 data items considered for B cohort – longitudinal weight

Variable name	Variable label
i1cnfsad2	16/17 - SC - SEIFA - Index of Relative Socio-Economic Advantage and Disadvantage - 2016 - SA2 - Score
i1cnfsad2d	16/17 - SC - SEIFA - Index of Relative Socio-Economic Advantage and Disadvantage - 2016 - SA2 - Score - Deciles - National
i1cnfser2	16/17 - SC - Home - SEIFA Economic Resources - 2016 - SA2 - Score
i1cnfser2d	16/17 - SC - Home - SEIFA Economic Resources - 2016 - SA2 - Deciles - National
i1p1cawi	16/17 - Parent 1 CAWI data present
hf03hp1	P1@14/15 - Age
hf03hp2	P2@14/15 - Age
hf11hm	M@14/15 - Main language spoken at home
hf11hp1	P1@14/15 - Main language spoken at home
hf11m1	14/15 - SC - Main language spoken at home
hfd08a1	14/15 - P1 - P1B CASI B1.1/B1.3+W1-7 - School completion
hfd08a2a	14/15 - P1 - P1B CASI B1.2/B1.3+W1-7 - Completed other post-secondary qualification
hfd08a3a	14/15 - P1 - P1B CASI B1.3+W1-7 - Highest qualification
hfd08m1	14/15 - M - P1B CASI B1.1/B1.3+W1-7 - School completion
i1pw30a1a	16/17 - P1 - P CAWI D1.1 - Employment status
i1pc82c7c	16/17 - SC CAWI I11.3 - Current level of achievement in studies
i1pw30b1a	16/17 - P2 -P CAWI D1.1 - Employment status
i1fd36c	16/17 - SC - SC CAWI A1.1 - Number of people living with SC
i1p2cawi	16/17 - Parent 2 CAWI data present
zf02hp1	P1@14/15 - Sex
zf09hp1	P1@14/15 - Country of birth
zf12hp1	P1@14/15 - Indigenous Status
i1datint	16/17 - SC - Date of interview
i1fd36c2	16/17 - SC - SC CAWI A1.2.2 - Live with step/half,sibling(s)

Table E3: Wave 1 data items considered for K cohort – cross-sectional weight

Variable name	Variable label
caangb	4/5—P1—Angry parenting (v3)
cahact	4/5—P1—Home activities index
ccnfsad	4/5—Home—SEIFA Advantage/Disadvantage
ccnfseo	4/5—Home—SEIFA Education & Occupation
ccnfser	4/5—Home—SEIFA Economic Resources
cf01cm	4/5—M@4/5—Present for wave
cf01m3	4/5—P2@W1—Present for wave
cf03m2	4/5—P1@W1—F2F A4—Age
cf03m3	4/5—P2@W1—F2F A4—Age
cf11cm	4/5—M@4/5—F2F A12—Main language spoken at home
cf11m1	4/5—SC—F2F A12—Main language spoken at home
cf11m2	4/5—P1@W1—F2F A12—Main language spoken at home
cfd08a1	4/5—P1—F2F H3—School completion
cfd08m1	4/5—M—F2F H3—School completion
cfd11m2	4/5—M—F2F H10—Proficiency in spoken English
cho04a3b	4/5—P1—F2F L4—Rent home
cho04a5	4/5—P1—F2F L5—Housing tenure
cho09a1a1	4/5—P1—F2F L11—Safe neighbourhood
cnpeople	4/5—No. people in household
cnsib	4/5—No. siblings of SC in household
cp1scd	4/5—Parent 1 self-completed data present
cp2	4/5—SC has 2 parents in the home
cp2scd	4/5—Parent 2 self-complete data present
zf02m2	P1@W1—F2F A3—Sex
zf09m2	P1@W1—F2F A10—Country of birth
zf12m1	SC—F2F A13—Indigenous status
zf12m2	P1@W1—F2F A13—Indigenous status
stratum	Stratum
zf02m1	SC - F2F A3 - Sex

Table E4: Wave 9C1 data items considered for K cohort – longitudinal weight

Variable name	Variable label
k1cnfsad2	20/21 - SC - SEIFA - Index of Relative Socio-Economic Advantage and Disadvantage - 2016 - SA2 - Score
k1cnfsad2d	20/21 - SC - SEIFA - Index of Relative Socio-Economic Advantage and Disadvantage - 2016 - SA2 - Score - Deciles - National
k1cnfser2	20/21 - SC - Home - SEIFA Economic Resources - 2016 - SA2 - Score
k1cnfser2d	20/21 - SC - Home - SEIFA Economic Resources - 2016 - SA2 - Deciles - National
k1p2cawi	20/21 - Parent 2 CAWI data present
jf03jp1	P1@18/19 - Age
jf03jp2	P2@18/19 - Age
jf11jm	M@18/19 - Main language spoken at home
jf11jp1	P1@18/19 -Main language spoken at home
jf11m1	18/19 - SC - Main language spoken at home
jfd08a3a	18/19 - P1 - P CATI_A1.2+W1 - 7 - Highest qualification
jfd08m3a	18/19 - M - P CATI_A1.2+W1 - 7 - Highest qualification
k1pw30b1a	20/21 - P2 - P CAWI D1.1 - Employment status
k1pw30a1a	20/21 - P1 - P CAWI D1.1 - Employment status
k1fd32a	20/21 - P1 - P CAWI A2.1 - No. people living with P1
k1fd36c	20/21 - SC - SC CAWI A1.1 - Number of people living with SC
k1fd36c1	20/21 - SC - SC CAWI A1.2.1 - Live with my parent(s) or step-parent(s)
zf02jp1	P1@18/19 - Sex
zf09jp1	P1@18/19 - Country of birth
zf12jp1	P1@18/19 - Indigenous Status
stratum	20/21 - Stratum
zf02m1	20/21 - SC - Sex

Appendix F: Distributional checks of non-response modelling

In order to validate the logistic regression non-response adjustment procedure, the estimated response propensities have been plotted below. There are also plots of the final sample weight under each model, where the approximate proportion of units at the caps can be observed.

B cohort – cross-sectional weight

Figure F1: Distribution of estimated response propensities – B cohort cross-sectional weight

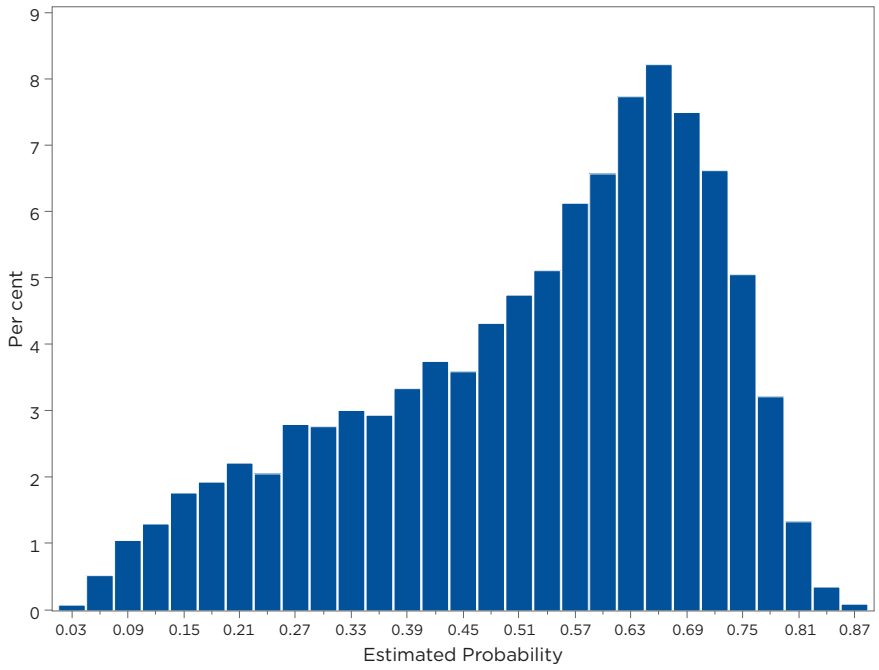

Table F1: Analysis variable: estimated probability – B cohort cross-sectional weight

Mean	Std Dev	Minimum	Maximum	Mode	Range	Sum	*N*
0.526339	0.18374	0.039734	0.87458	0.582075	0.83485	2,688	5,107

Figure F2: Distribution of final sample weight for Wave 9C2 – B cohort cross-sectional weight

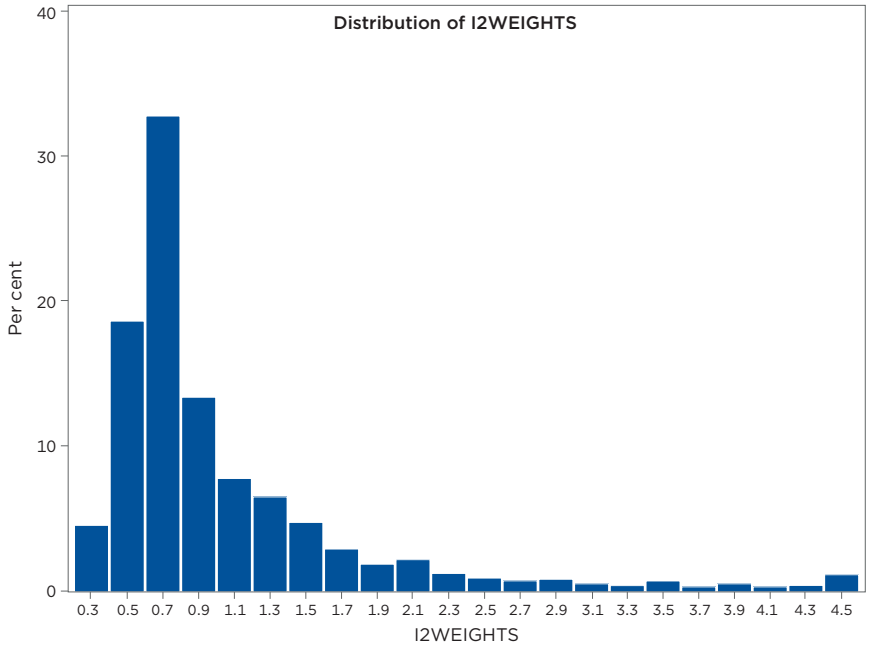

Table F2: Analysis variable: I2WEIGHTS– B cohort cross-sectional weight

Mean	Std Dev	Minimum	Maximum	Mode	Range	Sum	*N*
1	0.69731	0.28	4.5	4.5	4.22	2,688	2,688

B cohort – longitudinal weight

Figure F3: Distribution of estimated response propensities – B cohort longitudinal weight

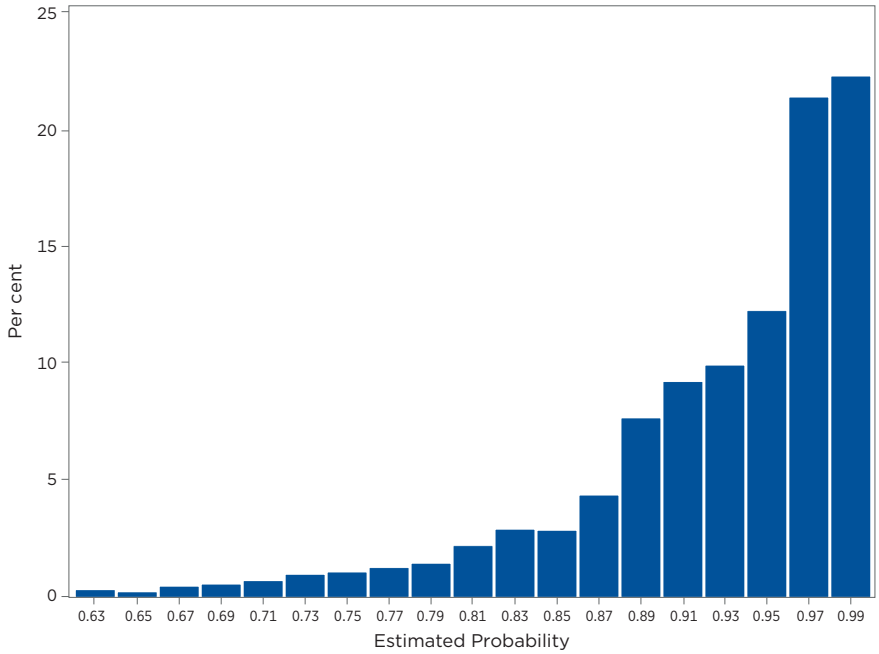

Table F3: Analysis variable: estimated probability – B cohort longitudinal weight

Mean	Std Dev	Minimum	Maximum	Mode	Range	Sum	*N*
0.92803	0.06691	0.6239	0.99781	0.89861	0.37392	1,612	1,737

Figure F4: Distribution of final sample weight for Wave 9C2 – B cohort longitudinal weight

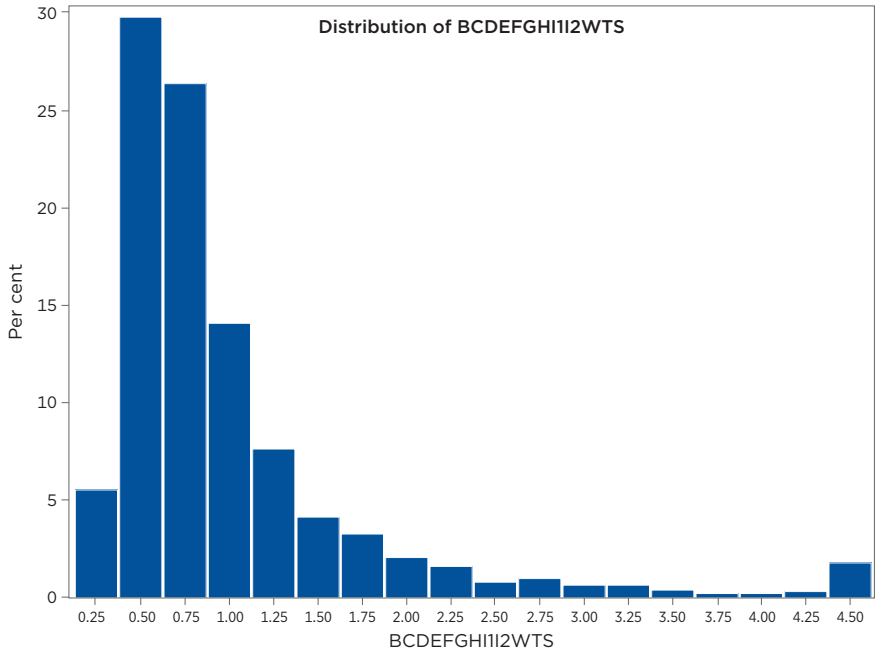

Table F4: Analysis variable: BCDEFGHI1I2WTS – B cohort longitudinal weight

Mean	Std Dev	Minimum	Maximum	Mode	Range	Sum	*N*
1	0.77456	0.28	4.5	4.5	4.22	1,612	1,612

K cohort – cross-sectional weight

Figure F5: Distribution of estimated response propensities – K cohort cross-sectional weight

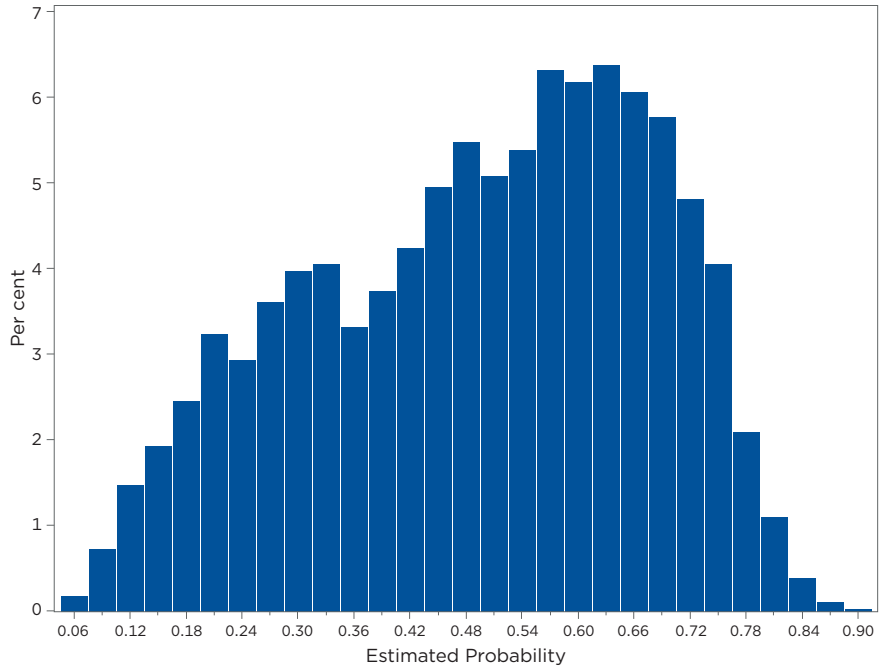

Table F5: Analysis variable: estimated probability – K cohort cross-sectional weight

Mean	Std Dev	Minimum	Maximum	Mode	Range	Sum	*N*
0.494281	0.18225	0.056989	0.889601	0.718736	0.83261	2,463	4,983

Figure F6: Distribution of final sample weight for Wave 9C2 – K cohort cross-sectional weight

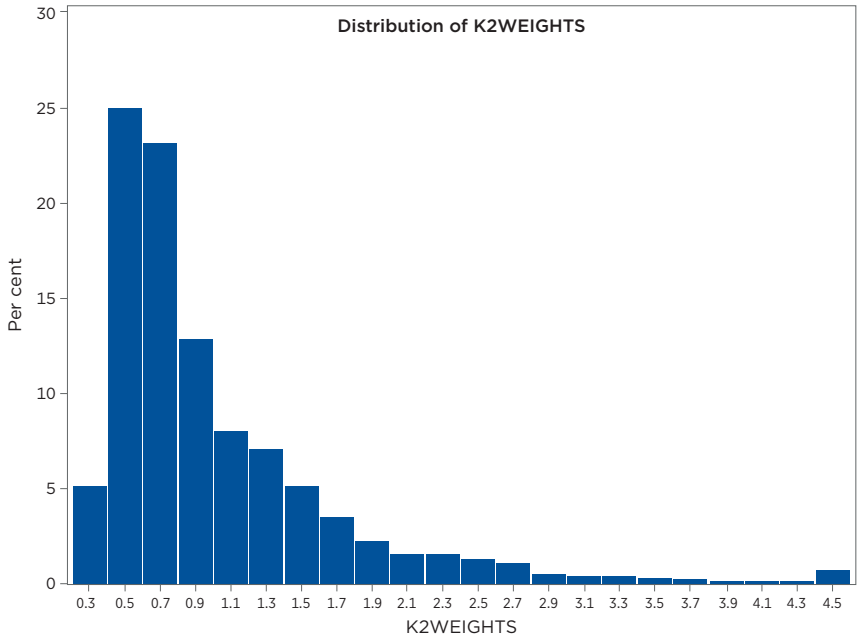

Table F6: Analysis variable: K2WEIGHTS – K cohort cross-sectional weight

Mean	Std Dev	Minimum	Maximum	Mode	Range	Sum	*N*
1	0.66979	0.28	4.5	0.28	4.22	2,463	2,463

K cohort – longitudinal weight

Figure F7: Distribution of estimated response propensities – K cohort longitudinal weight

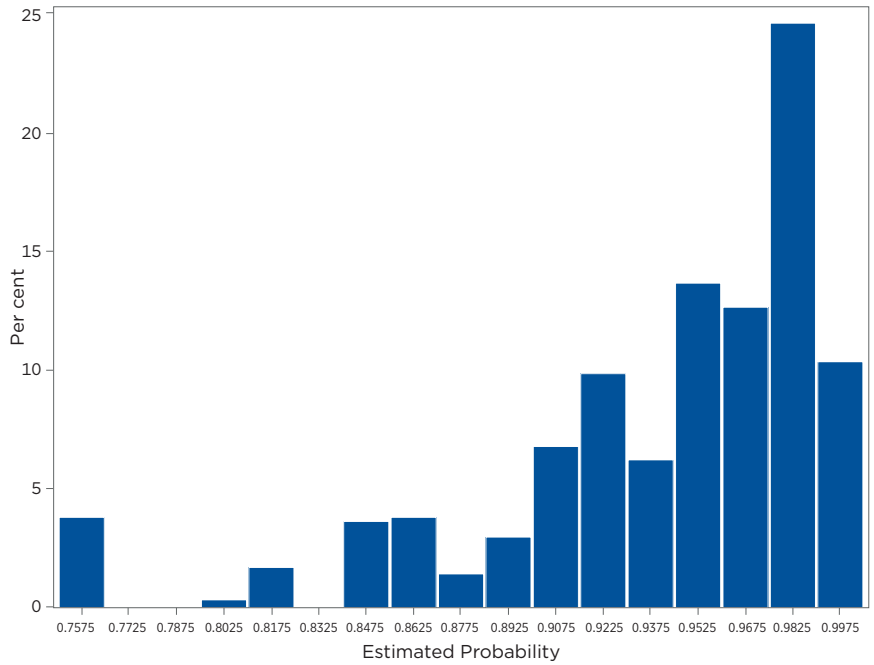

Table F7: Analysis variable: estimated probability – K cohort longitudinal weight

Mean	Std Dev	Minimum	Maximum	Mode	Range	Sum	*N*
0.94087	0.05514	0.75222	0.99998	0.95009	0.24775	1,432	1,522

Figure F8: Distribution of final sample weight for Wave 9C2 – K cohort longitudinal weight

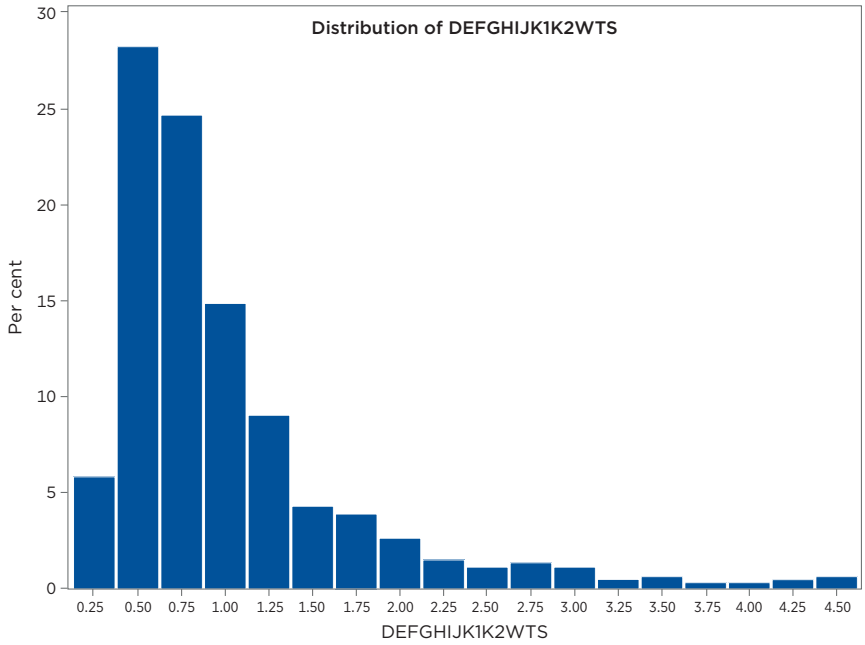

Table F8: Analysis variable: DEFGHIJK1K2WTS – K cohort longitudinal weight

Mean	Std Dev	Minimum	Maximum	Mode	Range	Sum	*N*
1	0.71315	0.28	4.5	0.28	4.22	1,432	1,432

Appendix G: Non-response to instruments

Table G1: Non-response to instruments

	Eligible	Responding	% Wave 1	Response rate %
	B cohort			
Wave 9C1 (issued sample = 3,849)				
SC CAWI	3,849	1,595	31.2	41.4
P1 CAWI	3,844	1,296	25.4	33.7
P2 CAWI	2,542	770	na	30.3
PLE CAWI	442	130	na	29.4
Wave 9C2 (issued sample = 3,716)				
SC CAWI	3,716	2,030	39.7	54.6
SC CATI*	2,827	198	na	7.0
P1 CAWI	3,710	1,810	35.4	48.8
P1 CATI*	2,985	389	na	13.0
	K cohort			
Wave 9C1 (issued sample = 3,809)				
SC CAWI	3,809	1,361	27.3	35.7
P1 CAWI (W7P1)	3,110	975	19.6	31.4
P2 CAWI (W7P2)	2,020	541	na	26.8
PLE CAWI (W7PLE)	413	110	na	26.6
Wave 9C2 (issued sample = 3,742)				
SC CAWI	3,742	1,763	35.4	47.1
SC CATI*	3,024	197	na	6.5
P1 CAWI	3,135	1,544	31.0	49.3
P1 CATI*	2,509	400	na	15.9

For appendix G, the issued sample numbers are for representations only, not for direct comparisons. This is due to a change in methodology for Wave 9C1.

*Respondents who had not completed their CAWI by 18 July, with a phone number on file, were included in this sample.

Instrument	Description
CAWI	Computer Assisted Web Interview
SC/P1/P2/PLE CATI	Study child/Parent 1/Parent 2/Parent living elsewhere Computer Assisted Telephone Interview
CAWSI	Computer Assisted Web Self Interview
EHC	Event History Calendar
CAI	Computer Assisted Interview
P1CASI	Parent 1 Computer Assisted Self Interview
P2SC	Parent 2 Self-Complete Questionnaire
Teach	Teacher Questionnaire
ACASI	Audio-Computer Assisted Self Interview
CSR	Child Self Report
TUD	Time Use Diary
MR	Matrix Reasoning
EXEC	Executive Functioning (CogState)
GJA	Rice Test of Grammatical Judgement
na	Not appropriate to compare with Wave 1

Appendix H: Joint distributions of response status and each predictor

These tables show how each variable in the response propensity model looks within the responding sample versus the non-responding households. The tables show how each categorical variable is distributed between the responding and non-responding households. Each tab finishes with a single table for all the continuous variables showing the mean and standard deviation of each variable in the responding and non-responding households.

Variable categories can be collapsed or re-parameterised as part of the weighting process. Small categories may have been collapsed during this process. Re-parameterisation is the re-labelling of modal categories. The odds ratios are calculated using the maximum category. Re-labelling the reference category to be the maximum makes the results easier to interpret. All the information needed about changes in categories is contained in the tables. Please do not compare these results against the Data Dictionary.

B cross-sectional

		Highest year of primary or secondary school completed by mother – afd08m1						
		Don't know, not asked, Refusal	Yr 11 or equivalent	Yr 10 or equivalent	Yr 9 or equivalent	Yr 8 or below, never attended school	Year 12 or equivalent	Total
Non-responding	**Freq**	6	327	523	132	72	1,359	2,419
	%	0.12	6.40	10.24	2.58	1.41	26.61	47.37
Responding	**Freq**	3	248	336	34	16	2,051	2,688
	%	0.06	4.86	6.58	0.67	0.31	40.16	52.63
Total	**Freq**	9	575	859	166	88	3,410	5,107
	%	0.18	11.26	16.82	3.25	1.72	66.77	100.00

		Renting home – aho04a3b			
		Refusal or don't know	Yes	No	Total
Non-responding	**Freq**	5	900	1,514	2,419
	%	0.10	17.62	29.65	47.37
Responding	**Freq**	2	573	2,113	2,688
	%	0.04	11.22	41.37	52.63
Total	**Freq**	7	1,473	3,627	5,107
	%	0.14	28.84	71.02	100.00

		Parent 1 Self-complete data added to file – ap1scd		
		No	Yes	Total
Non-responding	**Freq**	539	1,880	2,419
	%	10.55	36.81	47.37
Responding	**Freq**	227	2,461	2,688
	%	4.44	48.19	52.63
Total	**Freq**	766	4,341	5,107
	%	15.00	85.00	100.00

		Parent 2 Self-complete data added to file – ap2scd			
		Not asked	No	Yes	Total
Non-responding	**Freq**	355	593	1,471	2,419
	%	6.95	11.61	28.80	47.37
Responding	**Freq**	122	341	2,225	2,688
	%	2.39	6.68	43.57	52.63
Total	**Freq**	477	934	3,696	5,107
	%	9.34	18.29	72.37	100.00

		Parent 1 Country of birth – zf09m2		
		Elsewhere	Australia	Total
Non-responding	**Freq**	581	1,838	2,419
	%	11.38	35.99	47.37
Responding	**Freq**	530	2,158	2,688
	%	10.38	42.26	52.63
Total	**Freq**	1,111	3,996	5,107
	%	21.75	78.25	100.00

		Study child sex – zf02m1		
		Male	Female	Total
Non-responding	**Freq**	1,275	1,144	2,419
	%	24.97	22.40	47.37
Responding	**Freq**	1,333	1,355	2,688
	%	26.10	26.53	52.63
Total	**Freq**	2,608	2,499	5,107
	%	51.07	48.93	100.00

		Mother's main language spoken at home – af11am		
		Non-English	English	Total
Non-responding	**Freq**	442	1,977	2,419
	%	8.65	38.71	47.37
Responding	**Freq**	301	2,387	2,688
	%	5.89	46.74	52.63
Total	**Freq**	743	4,364	5,107
	%	14.55	85.45	100.00

	Variable	Label	Observations	Mean	SD	Minimum	Maximum
Non-responding	acnfseo	0/1 - Home - SEIFA Education & Occupation	2,419	985.14	75.41	783.2	1222.48
	af03m2	0/1 - P1@W1 - F2F A4 - Age	2,419	30.00	5.89	0	63
	af03m3	0/1 - P2@W1 - F2F A4 - Age	2,419	28.11	13.05	0	65
	ansib	0/1 - No. siblings of SC in household	2,419	1.05	1.13	0	8
Responding	acnfseo	0/1 - Home - SEIFA Education & Occupation	2,688	1,006.44	81.65	690.24	1,222.48
	af03m2	0/1 - P1@W1 - F2F A4 - Age	2,688	31.91	5.00	16	55
	af03m3	0/1 - P2@W1 - F2F A4 - Age	2,688	32.76	9.07	0	62
	ansib	0/1 - No. siblings of SC in household	2,688	0.88	1.00	0	10

B longitudinal

		Parent 1 completed CAWI – l1p1cawi		
		Not asked	Yes	Total
Non-responding	**Freq**	69	56	125
	%	3.97	3.22	7.20
Responding	**Freq**	507	1,105	1,612
	%	29.19	63.62	92.80
Total	**Freq**	576	1,161	1,737
	%	33.16	66.84	100.00

		Parent 2 completed CAWI – l1p2cawi		
		Not asked	Yes	Total
Non-responding	**Freq**	87	38	125
	%	5.01	2.19	7.20
Responding	**Freq**	928	684	1,612
	%	53.43	39.38	92.80
Total	**Freq**	1,015	722	1,737
	%	58.43	41.57	100.00

		Study child current level of achievement in studies – i1pc82c7c							
		Not asked	Missed	Very high	High	Average	Low	Very low	Total
Non-responding	**Freq**	73	1	1	10	29	5	6	125
	%	4.20	0.06	0.06	0.58	1.67	0.29	0.35	7.20
Responding	**Freq**	451	12	122	377	475	129	46	1,612
	%	25.96	0.69	7.02	21.70	27.35	7.43	2.65	92.80
Total	**Freq**	524	13	123	387	504	134	52	1,737
	%	30.17	0.75	7.08	22.28	29.02	7.71	2.99	100.00

	Variable	Label	Observations	Mean	SD	Minimum	Maximum
Non-responding	hf03hp1	Parent 1 age	125	44.97	5.46	34	60
Responding	hf03hp1	Parent 1 age	1,612	46.51	4.90	0	65

K cross-sectional

		Parent 1 Main language spoken at home – cf11m2		
		Non-English	English	Total
Non-responding	**Freq**	447	2,073	2,520
	%	8.97	41.60	50.57
Responding	**Freq**	330	2,133	2,463
	%	6.62	42.81	49.43
Total	**Freq**	777	4,206	4,983
	%	15.59	84.41	100.00

		Highest year of primary or secondary school completed by mother – cfd08m1								
		Not asked	Don't know or refusal	Year 11 or equivalent	Year 10 or equivalent	Year 9 or equivalent	Year 8 or below	Never attended school	Year 12 or equivalent	Total
Non-responding	**Freq**	27	5	388	639	151	90	5	1,215	2,520
	%	0.54	0.10	7.79	12.82	3.03	1.81	0.10	24.38	50.57
Responding	**Freq**	10	2	280	406	51	31	3	1,680	2,463
	%	0.20	0.04	5.62	8.15	1.02	0.62	0.06	33.71	49.43
Total	**Freq**	37	7	668	1,045	202	121	8	2,895	4,983
	%	0.74	0.14	13.41	20.97	4.05	2.43	0.16	58.10	100.00

		Renting home – cho04a3b			
		Refusal or don't know	Yes	No	Total
Non-responding	**Freq**	7	840	1,673	2,520
	%	0.14	16.86	33.57	50.57
Responding	**Freq**	2	489	1,972	2,463
	%	0.04	9.81	39.57	49.43
Total	**Freq**	9	1,329	3,645	4,983
	%	0.18	26.67	73.15	100.00

		Parent 1 Self-complete data added to file – cp1scd		
		No	Yes	Total
Non-responding	**Freq**	548	1,972	2,520
	%	11	39.57	50.57
Responding	**Freq**	206	2,257	2,463
	%	4.13	45.29	49.43
Total	**Freq**	754	4,229	4,983
	%	15.13	84.87	100.00

		Parent 2 Self-complete data added to file – cp2scd			
		Not asked	No	Yes	Total
Non-responding	**Freq**	488	611	1,421	2,520
	%	9.79	12.26	28.52	50.57
Responding	**Freq**	209	287	1,967	2,463
	%	4.19	5.76	39.47	49.43
Total	**Freq**	697	898	3,388	4,983
	%	13.99	18.02	67.99	100.00

		Stratum							
		NSW Met	NSW Xmet large	NSW Xmet small	Vic. Met large	Vic. Met small	Vic. Xmet large	Vic. Xmet small	Qld Met
Non-responding	**Freq**	478	268	29	423	11	138	51	216
	%	9.59	5.38	0.58	8.49	0.22	2.77	1.02	4.33
Responding	**Freq**	473	297	28	445	8	125	43	192
	%	9.49	5.96	0.56	8.93	0.16	2.51	0.86	3.85
Total	**Freq**	951	565	57	868	19	263	94	408
	%	19.08	11.34	1.14	17.42	0.38	5.28	1.89	8.19

		Stratum							
		Qld Xmet large	Qld Xmet small	SA Met large	SA Xmet large	SA Xmet small	WA Met large	WA Met small	WA Xmet large
Non-responding	**Freq**	290	42	138	40	9	169	6	46
	%	5.82	0.84	2.77	0.80	0.18	3.39	0.12	0.92
Responding	**Freq**	194	34	135	27	10	176	9	51
	%	3.89	0.68	2.71	0.54	0.20	3.53	0.18	1.02
Total	**Freq**	484	76	273	67	19	345	15	97
	%	9.71	1.53	5.48	1.34	0.38	6.92	0.30	1.95

		Stratum						
		WA Xmet small	Tas. Met	Tas. Xmet	NT Met	NT Xmet small and NT Xmet large	ACT	Total
Non-responding	**Freq**	28	21	42	19	21	35	2,520
	%	0.56	0.42	0.84	0.38	0.42	0.70	50.57
Responding	**Freq**	22	33	40	23	20	78	2,463
	%	0.44	0.66	0.80	0.46	0.40	1.57	49.43
Total	**Freq**	50	54	82	42	41	113	4,983
	%	1.00	1.08	1.65	0.84	0.82	2.27	100.00

		Study child sex - zf02m1		
		Male	Female	Total
Non-responding	**Freq**	1,348	1,172	2,520
	%	27.05	23.52	50.57
Responding	**Freq**	1,188	1,275	2,463
	%	23.84	25.59	49.43
Total	**Freq**	2,536	2,447	4,983
	%	50.89	49.11	100.00

	Variable	Label	Observations	Mean	SD	Minimum	Maximum
Non-responding	ccnfseo	4/5 - Home - SEIFA Education & Occupation	2,520	985.97	75.84	783.20	1221.12
	cf03m2	4/5 - P1@W1 - F2F A4 - Age	2,520	33.86	5.92	0	73
Responding	ccnfseo	4/5 - Home - SEIFA Education & Occupation	2,463	1,010.65	80.47	835.60	1,209.20
	cf03m2	4/5 - P1@W1 - F2F A4 - Age	2,463	35.62	4.94	0	65

K longitudinal

		Parent 1 Employment status – k1pw30a1a				
		Not asked	Missed	Full-time employee	Part-time employee	Self-employed
Non-responding	**Freq**	60	1	12	5	4
	%	3.94	0.07	0.79	0.33	0.26
Responding	**Freq**	565	12	357	277	90
	%	37.12	0.79	23.46	18.20	5.91
Total	**Freq**	625	13	369	282	94
	%	41.06	0.85	24.24	18.53	6.18

		Parent 1 Employment status – k1pw30a1a			
		Unpaid family worker	Unemployed	Not seeking employment	Total
Non-responding	**Freq**	1	1	6	90
	%	0.07	0.07	0.39	5.91
Responding	**Freq**	10	26	95	1,432
	%	0.66	1.71	6.24	94.09
Total	**Freq**	11	27	101	1,522
	%	0.72	1.77	6.64	100.00

		Study child sex – zf02m1		
		Male	Female	Total
Non-responding	**Freq**	56	34	90
	%	3.68	2.23	5.91
Responding	**Freq**	660	772	1,432
	%	43.36	50.72	94.09
Total	**Freq**	716	806	1,522
	%	47.04	52.96	100.00

	Variable	Label	Observations	Mean	SD	Minimum	Maximum
Non-responding	k1fd36c	20/21 - SC CAWI - Number of people living with SC	90	1.36	1.6	0	7
Responding	k1fd36c	20/21 - SC CAWI - Number of people living with SC	1,432	2.20	1.87	0	19