



**Longitudinal Study of Australian Children's
Child Health CheckPoint**

**Technical Paper 1: Weighting and Non-Response
December 2018**

Susan Ellul, Richard Hiscock, Fiona Mensah, Susan Clifford and
John Carlin



The Longitudinal Study of Australian Children

Abbreviations

ABS	Australian Bureau of Statistics
CheckPoint	Child Health CheckPoint
LSAC	Longitudinal Study of Australian Children
SEIFA	Socio-Economic Index for Areas

Acknowledgements and suggested citation

This paper draws on the work of Ryan Defina and Benedict Cusack in the development of weights for the Longitudinal Study of Australian Children's (LSAC) wave 5, described in LSAC Technical Paper No. 10: Wave 5 weighting and non-response (Cusack, 2013), as well as that of Amanda Norton and Kevin Monahan who developed the LSAC wave 6 cross sectional weights used in this paper and described in LSAC Technical Paper No. 15: Wave 6 weighting and non-response (Norton & Monahan, 2015).

Readers wishing to refer to this document should cite the following:

Susan Ellul, Richard Hiscock, Fiona Mensah, Susan Clifford and John Carlin. (2018). Longitudinal Study of Australian Children's Child Health CheckPoint Technical Paper 1: Weighting and Non-Response. Melbourne: Murdoch Children's Research Institute. doi: [10.25374/MCRI.5687593](https://doi.org/10.25374/MCRI.5687593)

About the authors

Susan Ellul is a Biostatistician/Research Assistant at the Clinical Epidemiology and Biostatistics Unit, Murdoch Children's Research Institute.

Dr Richard Hiscock is an anaesthetist (MBBS FANZCA) who completed the initial version of this work as a research project for his Master of Biostatistics degree.

Dr Fiona Mensah is a Biostatistician at the Clinical Epidemiology and Biostatistics Unit, Murdoch Children's Research Institute. She is also an Honorary Research Affiliate, Department of Paediatrics, University of Melbourne.

Dr Susan Clifford is a Postdoctoral Research Officer at the Centre for Community Child Health, Murdoch Children's Research Institute, and the Project Manager of LSAC's Child Health CheckPoint. She is also an Honorary Fellow of the Department of Paediatrics, University of Melbourne.

Professor John Carlin is Director of the Clinical Epidemiology and Biostatistics Unit at the Murdoch Children's Research Institute and University of Melbourne Department of Paediatrics. He is also Professorial Fellow, Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, University of Melbourne.

Contact information

For more information, contact:

Child Health CheckPoint team
Murdoch Children's Research Institute
50 Flemington Road PARKVILLE VIC 3052
Website: checkpoint-lsac.mcri.edu.au
Email: lsac.childhealthcheckpoint@mcri.edu.au
Phone: +61 3 9936 6464

Table of Contents

Abbreviations	2
Acknowledgements and suggested citation	3
About the authors	3
Contact information.....	3
1. Introduction.....	5
2. CheckPoint weighting method.....	7
2.1. Initial weights	7
2.2. Modelling the response propensity of subsample participation	7
2.2.1 Covariate selection for logistic regression.....	8
2.2.2 Variable selection and shrinkage of estimated coefficients	9
2.3. Stratum weight adjustment	9
2.4. Weight capping.....	10
3. Conclusion	10
4. References.....	11
Appendices	12
Appendix A: Description of CheckPoint survey weights	12
Appendix B: Covariates considered for Child Health CheckPoint response propensity models	13
Appendix C: Included variables and odds ratio estimates for Child Health CheckPoint response propensity models	19
Appendix D: Distribution of the five final sets of survey weights.....	29

1. Introduction

The Growing Up in Australia’s Child Health CheckPoint study is a cross-sectional study of Australian children in the B cohort of the Longitudinal Study of Australian Children (LSAC). The CheckPoint study was conducted between the data collection for waves 6 and 7 of LSAC, when the children in the cohort were aged between 11-12 years. CheckPoint is a comprehensive, one-off physical health and biomarker module that aimed to collect data on physical health measurements (cardiovascular, respiratory, bone density, vision, hearing), biological samples (DNA and biomarkers, tissue samples) and digital information (images and activity monitors; Wake et al., 2014).

LSAC recruited two cohorts of children: the B cohort that consisted of children aged less than one year old, born between March 2003 and February 2004; and the K cohort of children who were aged between four to five years old and born between March 1999 and February 2000. The first wave of data collection took place in 2004, with subsequent main waves of data collection conducted every two years. In 2005, 2007, 2009 and 2011 parents were also sent a mail-based survey. LSAC Technical Paper 1 (Soloff, Lawrence, & Johnstone, 2005) provides more information on the design of LSAC and Technical Paper 3 provides details of the LSAC wave 1 sample (Soloff, Lawrence, Mission, & Johnstone, 2006). The initially targeted group for the B-cohort, including non-contacts, was 8,921 (approximately 4% of the reference population) and the recruited sample size at wave 1 was 5,107 (approximately 2% of the reference population).

Wave 6 of the LSAC was conducted in 2014 with the B cohort aged between 10 – 11 years old and there were 3,764 children who participated in the survey at that time (73.7% of the original wave 1 cohort). The number of children participating from the B cohort decreased with each assessment wave 2-6 and decreased further for the CheckPoint study (Table 1). The number of children who participated in CheckPoint was 1,874, which represented 49.8% and 36.7% of the LSAC B cohort at waves 6 and 1 respectively.

Table 1. Response characteristics for the LSAC B cohort and CheckPoint study

Year	2004	2006	2008	2010	2012	2014	2015
Wave	1	2	3	4	5	6	CheckPoint
Cross-sectional response	5,107	4,606	4,386	4,242	4,085	3,764	1,874
Attrition Rate (%)*		9.8	14.1	16.9	20.0	26.3	63.3

* Cumulative attrition rate – non-responders at current wave / 5,107

The CheckPoint sampling frame was therefore all 3,764 study children who both responded to and took part in the LSAC wave 6 assessment. Letters of invitation, detailing the CheckPoint study aims and schedule, were sent to the adult registered as the primary parent of the study child (cohort member). Given the nature of assessments, it was decided that researcher home visits (used in main LSAC waves) were not the ideal data collection module for CheckPoint. Instead, Assessment Centres would be set up in major cities, supplemented by a mobile assessment team who set up Mini-Assessment Centres in regional towns. Shortened home visits were offered to families who could not attend an Assessment Centre. The burden for responders (children and attending parent (P1)) was therefore considerable, consisting of both time (with assessment time approximately 3.5 hours plus travel time to centres) and the inconvenience of assessments that occurred during the school and working week. Many factors were likely to predict participation in LSAC and these factors were also important in predicting CheckPoint participation. However, given the considerable additional burden associated with CheckPoint compared to LSAC participation, several factors beyond those that were predictive of LSAC completion were important predictors for CheckPoint participation. Where attendance at an Assessment Centre was not possible, Home Visits and Mini Assessment Centre Visits were offered as an alternative, but for these types of visits CheckPoint assessment was only available on a subset of measures. Recruitment commenced in late 2014 with final assessments conducted in March 2016.

This paper details the methods used to develop cross-sectional survey weights for CheckPoint. Five sets of survey weights have been produced for subsamples defined by CheckPoint participation. Those who participated in CheckPoint at all were likely to differ from those who did not participate in CheckPoint and hence survey weights for those who participated in CheckPoint were produced. In addition, those who attended a Main Assessment Centre rather than attending a Mini Assessment Centre or having a Home Visit had the opportunity to undertake the full protocol of measures and therefore a set of weights for those who attended a Main Assessment Centre only were produced. A set of survey weights has also been developed for those who attended either a Main Assessment Centre or Mini Assessment Centre because despite those attending Mini Assessment Centres not having the opportunity to undertake the full range of measures, those attending such centres shared the opportunity to undertake a large number of measures compared to those who had Home Visits. Lastly, those who provided a blood sample were likely to differ from those who did not and therefore survey weights were produced for those who provided a blood sample. As the study child and attending parent may not have necessarily both provided blood samples, two sets of survey weights have been produced; one for the study child who provided blood samples and another for the attending parent who provided a blood sample. A summary of the survey weights produced for CheckPoint are provided in Appendix A.

The weights have been developed for use by data users who are using data from the CheckPoint study and are recommended to be used in analyses targeting population descriptive quantities such as mean values, in order to make less biased population-based inferences.

2. CheckPoint weighting method

The method used to generate weights was similar to that used to create the LSAC waves 5 and 6 final sample weights (Norton & Monahan, 2015). There were three main stages for the development of the CheckPoint weights. The first was modelling the response propensity of CheckPoint subsample participation based on factors known and deemed as important for LSAC wave 6 responders and calculating inverse probability weights from the estimated response propensities. The second step was combining the inverse probability weights with the weights provided for LSAC wave 6 in order to adjust for response propensity relative to the original target sample. The final step was the calibration of the adjusted weights to reflect the composition of the original B cohort at wave 1. As part of the final step, weights were capped.

2.1. *Initial weights*

In longitudinal surveys the final sample weights for the previous wave are used as the initial weights in the development of weights for the next wave. Hence, the final sample cross-sectional LSAC wave 6 weights were used as the initial weights for development of the CheckPoint cross-sectional weights.

2.2. *Modelling the response propensity of subsample participation*

The aim was to obtain predicted probabilities of participation for each member of each CheckPoint subsample based upon important factors that were known for LSAC wave 6 participants. Factors deemed to be predictors of participation in each CheckPoint subsample were included in logistic regression models. The resulting response probabilities were inverted to produce inverse probability weights which were then used in the calculation of the final CheckPoint weights. For each CheckPoint subsample (overall participation, attendance at Main Assessment Centre, attendance at Main Assessment Centre or Mini Assessment Centre, study child blood sample, and attending parent blood sample), a separate logistic regression model was used in order to produce five separate sets of cross-sectional survey weights.

2.2.1 Covariate selection for logistic regression

There were over 4,000 measured covariates in the LSAC wave 6 study of 3,764 children. Given that it was not reasonable to include all of these covariates in the response propensity models, potential covariates were chosen based on substantive judgements of the likelihood of them being important predictors of appearing in each CheckPoint subsample, while restricting to covariates for which no more than 3% of values were missing.

To maintain consistency with potential covariates that were considered in the generation of the LSAC wave 6 B cohort weights, where possible, equivalent LSAC variables from wave 6 were initially considered as potential covariates for the generation of the Child Health CheckPoint cross-sectional weights.

Simple imputation techniques were used to “fill in” missing values of covariates that were used in the process of generating weights, in order to reduce complexity. Because such methods may introduce bias if used with large amounts of missing data (Harrell, 2015), variables that had more than 3% missing values were not considered further. As a result of the decision to exclude variables with more than 3% missing values, a small number of variables equivalent to those used for LSAC wave 6 weighting models were omitted from the set of potential covariates. For covariates with fewer than 3% missing values, imputation was undertaken using the median value for continuous and modal values for categorical variables.

For weights applicable to (any) **participation** (those who took part at all in Child Health CheckPoint), an additional 47 covariates were added to the potential covariate list because it was felt that these were the variables most likely to be predictors of participation. Examples of variables that were deemed likely to be predictors of participation in Child Health CheckPoint were those related to disability, financial measures and household commitments. Variables deemed not likely to be predictors of participation in Child Health CheckPoint included those related to dietary preference, allergies, sleeping habits, puberty details and relationship descriptors.

For weights applicable to those who attended a **Main Assessment Centre** (not Home Visit or Mini Assessment Centre) or **Main Assessment Centre/Mini Assessment Centre** (not Home Visit), the same set of potential covariates were used as for **participation**. It was felt that the same set of potential covariates would suffice as it would likely include predictors for whether a participant attended a Main Assessment Centre or attended a Mini Assessment Centre or Home Visit. The addition of further variables would likely only increase complexity without making any substantial difference.

For weights applicable to the study child and attending **parents** (adults) who provided a **blood sample**, two additional variables related to worries and fears were added to the potential covariate list because it was felt that they were additional potential predictors of whether a participant would provide a blood sample.

In summary, potential covariates in the response probability models were selected based on having no more than 3% missing values (to avoid complex imputation techniques) and compatibility with potential covariates for LSAC wave 6 weighting (with the above limitation). In addition, other variables were included as they were considered likely to be predictors of CheckPoint participation. For weighting based on blood sampling, two additional variables were included in the potential covariate list because it seemed plausible that worries and fears could be predictors of whether a blood sample was provided.

As a result, 72 covariates were initially considered for response propensity models generated for the subsamples relating to **overall participation** and whether they attended a **Main**

Assessment Centre or **Main Assessment Centre/Mini Assessment Centre**. There were 74 potential covariates considered for both sets of **blood sampling** weights. The list of potential covariates considered for each type of survey weight generated can be found in Appendix B.

2.2.2 Variable selection and shrinkage of estimated coefficients

For each type of weight, backward stepwise regression was used to further reduce the number of selected covariates from those listed in Appendix B. The method used was consistent with the method that has been used in LSAC weighting since wave 4 (Norton & Monahan, 2015). In the modelling, the p-value criteria used for inclusion and exclusion were 0.10 and 0.05 respectively and the p-value was determined by using the likelihood ratio χ^2 statistic between adjacent nested models with the appropriate degrees of freedom.

Shrinkage of regression coefficients was undertaken to adjust for over-fitting and improve calibration (Steyerberg, Eijkemans, & Habbema, 2001). To do this, a global shrinkage factor was applied to estimated regression coefficients prior to predicting response probabilities (Van Houwelingen & Le Cessie, 1990).

Following the selection of variables and application of the shrinkage factor, predicted probabilities of belonging to a relevant CheckPoint subsample were calculated from the corresponding fitted response propensity model.

Following this approach, there were 22, 19, 16, 20 and 21 variables found to be predictive of overall CheckPoint participation, Main Assessment Centre attendance, Main Assessment Centre/Mini Assessment Centre attendance, study child blood sample and attending parent blood sample respectively. The adjusted coefficients for the corresponding models with global shrinkage are presented as odds ratios in Appendix C.

2.3. Stratum weight adjustment

The inverse probability weights were then adjusted back to non-calibrated weights at wave 1. To do this, inverse probability weights were generated by taking the reciprocal of the estimated response probabilities and multiplying by the sample cross-sectional LSAC wave 6 weights (Soloff et al., 2006). Calibration, or post-stratification adjustment, was then used to align the survey weights to the composition within each stratum at LSAC wave 1 whilst ensuring the survey weights sum to the relevant CheckPoint sample size.

To be consistent with methods used for LSAC weighting for waves 5 and 6 (Norton & Monahan, 2015), a contingency table based on the cross classification of the cohort by state, region and postcode population size was used to divide the LSAC wave 1 cohort and the CheckPoint participants into mutually exclusive groups. The weighted cell counts for the CheckPoint cohort were then recalibrated to the population counts applicable to LSAC wave 1. The calibration was performed using an iterative proportional fitting (raking) algorithm which attempts to match marginal stratum totals whilst constrained by total sample size (Kolenikov, 2014). For the weight adjustment, some collapsing of strata was required for the Main Assessment Centre CheckPoint weights in order to obtain calibration whilst satisfying weight caps.

2.4. *Weight capping*

The weighting process can produce extreme weights and these large weights may exert undue influence on the value of calculated weighted survey items, thus considerably inflating the variance of weighting-based estimates of population parameters. In the development of LSAC wave 1-6 weights, weights were top and bottom coded to constrain weights to be greater than 0.33 and less than 2.5 (Norton & Monahan, 2015). Whilst these settings may be appropriate in settings with low attrition rates, in the CheckPoint study there was a higher attrition rate (36.7% of the original LSAC cohort participated). With attrition, the number of weights at the weight caps increases as has been observed with successive waves of LSAC. In order to ensure that the CheckPoint weights were comparable in distribution, particularly to ensure comparability in the proportion of weights at lower and upper caps, a lower trimming of 0.25 and an upper trimming of 3.5 was used for the CheckPoint survey weights.

The distributions of the final cross sectional survey weights are presented in appendix D.

3. Conclusion

Five sets of CheckPoint survey weights have been developed for use by data users using data from the CheckPoint study. The weights were produced for each subsample defined by level of CheckPoint participation: participation itself (at any level), attendance at a Main Assessment Centre, attendance at a Main Assessment Centre/Mini Assessment Centre, study children who produced a blood sample and attending parents who produced a blood sample. For each subsample, survey cross-sectional weights have been developed.

4. References

- Cusack, B. D., Ryan. (2013). LSAC Technical Paper No. 10. Wave 5 weighting and non response.
- Harrell, F. (2015). *Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis*: Springer.
- Kolenikov, S. (2014). Calibrating survey data using iterative proportional fitting (raking). *Stata Journal*, 14(1), 22-59.
- Norton, A., & Monahan, K. (2015). LSAC Technical Paper No. 15. Wave 6 Weighting and Non-Response. *ABS*.
- Soloff, C., Lawrence, D., & Johnstone, R. (2005). LSAC Technical Paper No. 1 Sample Design. *ABS*.
- Soloff, C., Lawrence, D., Mission, S., & Johnstone, R. (2006). LSAC Technical paper No. 3. Wave 1 weighting and non-response. *ABS*.
- Steyerberg, E. W., Eijkemans, M. J. C., & Habbema, J. D. F. (2001). Application of Shrinkage Techniques in Logistic Regression Analysis: A Case Study. *Statistica Neerlandica*, 55(1), 76-88. doi:10.1111/1467-9574.00157
- Van Houwelingen, J. C., & Le Cessie, S. (1990). Predictive value of statistical models. *Statistics in Medicine*, 9(11), 1303-1325. doi:10.1002/sim.4780091109
- Wake, M., Clifford, S., York, E., Mensah, F., Gold, L., Burgner, D., & Davies, S. (2014). Introducing growing up in Australia's child health check point: A physical health and biomarkers module for the longitudinal study of Australian children. *Family Matters*(95), 15.

Appendices

Appendix A: Description of CheckPoint survey weights

Variable name	CheckPoint subsample	Type/To be used for	Multiplier to use to obtain population weights*
<i>fweightscp</i>	Overall CheckPoint participants	Cross sectional survey weight to be used for all study children or all attending parents ¹ who participated in CheckPoint (n=1874)	129.68
<i>fweightsmn</i>	Main Assessment Centre CheckPoint participants	Cross sectional survey weight to be used for all study children or all attending parents who attended a CheckPoint Main Assessment Centre (not those who had a Home Visit or attended a Mini Assessment Centre) (n=1356)	179.22
<i>fweightsac</i>	Main Assessment Centre AND Mini Assessment Centre CheckPoint participants	Cross-sectional survey weight to be used for all study children or all attending parents who attended a CheckPoint Main Assessment Centre or Mini Assessment Centre- (not those who had a Home Visit) (n=1509). Note: if a measure was only available at the Main Assessment Centre and not the Mini Assessment Centre then the Main Assessment Centre weights should be used.	161.05
<i>fcweightsb</i>	Study child CheckPoint participants who provided a blood sample	Cross sectional survey weight to be used for study child CheckPoint participants who provided a blood sample (n=1237) or for pairs of study children and attending parents who both provided a blood sample (n=1200)	196.46
<i>faweightsb</i>	Attending parents of study children participating in Checkpoint who provided a blood sample	Cross sectional survey weight to be used for attending parents of study children participating in CheckPoint who provided a blood sample (n=1373)	177.00

¹In all cases parents also includes any attending adults who participated in CheckPoint who are not parents.

*If needed, population cross-sectional weights can be obtained by upscaling survey cross-sectional weights using an adjustment factor. The multiplier is the Australian Bureau of Statistics (ABS) estimated resident population counts of children aged 0 years at end of March 2004 (243,026) divided by the relevant CheckPoint subsample size.

Appendix B: Covariates considered for Child Health CheckPoint response propensity models

Table B1: Covariates considered for Child Health CheckPoint participation, attendance at Main Assessment Centre and attendance at Main Assessment Centre/Mini Assessment Centre survey weights

Number	Variable name	Variable label
(1)	faanga	10/11 - p1 - angry parenting (v2)
(2)	fairc	10/11 - p1 - parent 1 inductive reasoning scale (v3)
(3)	fak6	10/11 - p1 - k-6 depression scale
(4)	faparmonb	10/11 - p1 - parental monitoring version2
(5)	fnlivec1	10/11 - neighbourhood liveability (v3)
(6)	fpc09	10/11 - care at any time - total hours per week
(7)	fpedspse	10/11 - peds psychosocial health summary (v5)
(8)	fsatii	10/11 - sati introversion
(9)	fsatip	10/11 - sati persistence
(10)	fsatir	10/11 - sati reactivity
(11)	fbfath	10/11 - biological father in the home
(12)	fbmoth	10/11 - biological mother in the home
(13)	fdisinh	10/11 - someone in home has a disability
(14)	fdisop1	10/11 - someone in home other than p1 has a disability
(15)	ffn02a1	10/11 - p1 - f2f o1.1.1 - income - wages or salary
(16)	ffn02a2	10/11 - p1 - f2f o1.1.2 - income - business
(17)	ffn02a5	10/11 - p1 - f2f o1.1.3 - income - government payment
(18)	ffd09a1a	10/11 - p1 - f2f a1.4 - currently studying
(19)	fhe09a	10/11 - f2f m8.1- extra curricular - any
(20)	fhe09a1a	10/11 - f2f m8.1.1- extra curricular - community group
(21)	fhe09a2a	10/11 - f2f m8.1.2 - extra curricular - team sport
(22)	fhe09a3a	10/11 - f2f m8.1.3- extra curricular - individual sport
(23)	fhe09a4a	10/11 - f2f m8.1.4 - extra curricular - art etc.
(24)	fhe09a5a	10/11 - f2f m8.1.5- extra curricular - academic classes
(25)	fhe09a6a	10/11 - f2f m8.1.6 - extra curricular - new skill classes
(26)	fho04a3b	10/11 - p1 - f2f p1.6.2 - rent home

Number	Variable name	Variable label
(27)	fnewpart	10/11 - new partner of p1 in the home
(28)	fpc64x3	10/11 - p1 - f2f k1.2.3 - before - school (last year) - program at sc's school
(29)	fpc64x4	10/11 - p1 - f2f k1.2.4 - before - school (last year) - program - another school
(30)	fpc64x5	10/11 - p1 - f2f k1.2.5 - before - school (last year) - care centre not at school
(31)	fpc65x3	10/11 - p1 - f2f k1.4.3 - after - school (last year) - program at sc's school
(32)	fpc65x4	10/11 - p1 - f2f k1.4.4 - after - school (last year) - program at another school
(33)	fpc65x5	10/11 - p1 - f2f k1.4.5 - after - school (last year) - care centre not at a school
(34)	fpc65x6	10/11 - p1 - f2f k1.4.6 - after - school (last year) - family day care
(35)	fpc65x9	10/11 - p1 - f2f k1.4.9 - after - school (last year) - parent lives/lived elsewhere
(36)	fsc12a1l	10/11 - p1 - f2f j14.1.6 - used for sc - speech therapy
(37)	fsc12a1w	10/11 - p1 - f2f j14.1.5- used for sc - disability services
(38)	fsc12a1z	10/11 - p1 - f2f j14.1 - used services for sc in last year
(39)	fsc12a2z	10/11 - p1 - f2f j14.2 - need services for sc in last year
(40)	fsc13a1zz	10/11 - p1 - f2f j14.3 - used for family in last 12 months
(41)	fsc19m	10/11 - m - help person with disability living elsewhere
(42)	fsc19a	10/11 - p1 - help person with disability living elsewhere
(43)	zf02m1	10/11 - sc - sex
(44)	fregion	10/11 - region of residence
(45)	fabsra	10/11 - remoteness area (abs)
(46)	fho04a1	10/11 - p1 - f2f p1.6-8 - home ownership
(47)	fmemp	10/11 - m - employment status
(48)	fpa14a2	10/11 - p1 - f2f b8.2+w2-5 - parents separated
(49)	fpc37a2	10/11 - school - f2f c1.2 - school type
(50)	fremote	10/11 - remoteness area classification (aria)
(51)	fstate	10/11 - state of residence
(52)	foohactb	10/11 - out of home activities index (v2)
(53)	fcnfsad2	10/11 - seifa - index of relative socio-econ advantage & disadvantage – 2011-SA2-score

Number	Variable name	Variable label
(54)	fcnfsad2d	10/11 - seifa - index of relative socio-econ advantage & disadvantage – 2011-SA2-deciles-national
(55)	fcnfser2	10/11 - home - seifa economic resources - 2011 - score
(56)	fcnfser2d	10/11 - home - seifa economic resources - 2011 - deciles - national
(57)	ff03fp1	p1@10/11 - age
(58)	ff11fm	m@10/11 - language other than english spoken at home
(59)	ff11fp1	p1@10/11 - language other than english spoken at home
(60)	ff11m1	10/11 - sc - main language spoken at home
(61)	ffd08a1	10/11 - p1 - f2f a1.1/a1.3+w1-5 - school completion
(62)	ffd08a2a	10/11 - p1 - f2f a1.2/a1.3+w1-5 - comp other post-sec qualification
(63)	ffd08m1	10/11 - m - f2f a1.1/a1.3+w1-5 - school completion
(64)	faemp	10/11 - p1 - employment status
(65)	fho04a5	10/11 - p1 - f2f p1.6-8 - housing tenure
(66)	fnpeople	10/11 - no. people in household
(67)	fnsib	10/11 - no. siblings of sc in household
(68)	fp2	10/11 - sc has 2 parents in the home
(69)	zf02fp1	p1@10/11 - sex
(70)	zf09fp1	p1@10/11 - country of birth
(71)	zf12fp1	p1@10/11 - indigenous status
(72)	f1c08a3a	10/11 - p1 - f2f c7.4 - overall school achievement

ABS: Australian Bureau of Statistics; SEIFA: Socio-Economic Index for Areas.

Table B2: Covariates considered for Child Health CheckPoint study child and attending parent blood sample survey weights

Number	Variable name	Variable label
(1)	faanga	10/11 - p1 - angry parenting (v2)
(2)	fairc	10/11 - p1 - parent 1 inductive reasoning scale (v3)
(3)	fak6	10/11 - p1 - k-6 depression scale
(4)	faparmonb	10/11 - p1 - parental monitoring version2
(5)	fnlivec1	10/11 - neighbourhood liveability (v3)
(6)	fpc09	10/11 - care at any time - total hours per week
(7)	fpedspse	10/11 - peds psychosocial health summary (v5)
(8)	fsatii	10/11 - sati introversion
(9)	fsatip	10/11 - sati persistence
(10)	fsatir	10/11 - sati reactivity
(11)	fbfath	10/11 - biological father in the home
(12)	fbmoth	10/11 - biological mother in the home
(13)	fdisinh	10/11 - someone in home has a disability
(14)	fdisop1	10/11 - someone in home other than p1 has a disability
(15)	ffn02a1	10/11 - p1 - f2f o1.1.1 - income - wages or salary
(16)	ffn02a2	10/11 - p1 - f2f o1.1.2 - income - business
(17)	ffn02a5	10/11 - p1 - f2f o1.1.3 - income - government payment
(18)	ffd09a1a	10/11 - p1 - f2f a1.4 - currently studying
(19)	fhe09a	10/11 - f2f m8.1- extra curricular - any
(20)	fhe09a1a	10/11 - f2f m8.1.1- extra curricular - community group
(21)	fhe09a2a	10/11 - f2f m8.1.2 - extra curricular - team sport
(22)	fhe09a3a	10/11 - f2f m8.1.3- extra curricular - individual sport
(23)	fhe09a4a	10/11 - f2f m8.1.4 - extra curricular - art etc.
(24)	fhe09a5a	10/11 - f2f m8.1.5- extra curricular - academic classes
(25)	fhe09a6a	10/11 - f2f m8.1.6 - extra curricular - new skill classes

Number	Variable name	Variable label
(26)	fho04a3b	10/11 - p1 - f2f p1.6.2 - rent home
(27)	fnewpart	10/11 - new partner of p1 in the home
(28)	fpc64x3	10/11 - p1 - f2f k1.2.3 - before - school (last year) - program at sc's school
(29)	fpc64x4	10/11 - p1 - f2f k1.2.4 - before - school (last year) - program - another school
(30)	fpc64x5	10/11 - p1 - f2f k1.2.5 - before - school (last year) - care centre not at school
(31)	fpc65x3	10/11 - p1 - f2f k1.4.3 - after - school (last year) - program at sc's school
(32)	fpc65x4	10/11 - p1 - f2f k1.4.4 - after - school (last year) - program at another school
(33)	fpc65x5	10/11 - p1 - f2f k1.4.5 - after - school (last year) - care centre not at a school
(34)	fpc65x6	10/11 - p1 - f2f k1.4.6 - after - school (last year) - family day care
(35)	fpc65x9	10/11 - p1 - f2f k1.4.9 - after - school (last year) - parent lives/lived elsewhere
(36)	fsc12a1l	10/11 - p1 - f2f j14.1.6 - used for sc - speech therapy
(37)	fsc12a1w	10/11 - p1 - f2f j14.1.5- used for sc - disability services
(38)	fsc12a1z	10/11 - p1 - f2f j14.1 - used services for sc in last year
(39)	fsc12a2z	10/11 - p1 - f2f j14.2 - need services for sc in last year
(40)	fsc13a1zz	10/11 - p1 - f2f j14.3 - used for family in last 12 months
(41)	fsc19m	10/11 - m - help person with disability living elsewhere
(42)	fsc19a	10/11 - p1 - help person with disability living elsewhere
(43)	zf02m1	10/11 - sc - sex
(44)	fregion	10/11 - region of residence
(45)	fabsra	10/11 - remoteness area (abs)
(46)	fho04a1	10/11 - p1 - f2f p1.6-8 - home ownership
(47)	fmemp	10/11 - m - employment status
(48)	fpa14a2	10/11 - p1 - f2f b8.2+w2-5 - parents separated
(49)	fpc37a2	10/11 - school - f2f c1.2 - school type
(50)	fremote	10/11 - remoteness area classification (aria)
(51)	fstate	10/11 - state of residence
(52)	fohactb	10/11 - out of home activities index (v2)

Number	Variable name	Variable label
(53)	fcnfsad2	10/11 - seifa - index of relative socio-econ advantage & disadvantage - 2011-SA2-score
(54)	fcnfsad2d	10/11 - seifa - index of relative socio-econ advantage & disadvantage - 2011-SA2-deciles-national
(55)	fcnfser2	10/11 - home - seifa economic resources - 2011 - score
(56)	fcnfser2d	10/11 - home - seifa economic resources - 2011 - deciles - national
(57)	ff03fp1	p1@10/11 - age
(58)	ff11fm	m@10/11 - language other than english spoken at home
(59)	ff11fp1	p1@10/11 - language other than english spoken at home
(60)	ff11m1	10/11 - sc - main language spoken at home
(61)	ffd08a1	10/11 - p1 - f2f a1.1/a1.3+w1-5 - school completion
(62)	ffd08a2a	10/11 - p1 - f2f a1.2/a1.3+w1-5 - comp other post-sec qualification
(63)	ffd08m1	10/11 - m - f2f a1.1/a1.3+w1-5 - school completion
(64)	faemp	10/11 - p1 - employment status
(65)	fho04a5	10/11 - p1 - f2f p1.6-8 - housing tenure
(66)	fnpeople	10/11 - no. people in household
(67)	fnsib	10/11 - no. siblings of sc in household
(68)	fp2	10/11 - sc has 2 parents in the home
(69)	zf02fp1	p1@10/11 - sex
(70)	zf09fp1	p1@10/11 - country of birth
(71)	zf12fp1	p1@10/11 - indigenous status
(72)	flc08a3a	10/11 - p1 - f2f c7.4 - overall school achievement
(73)	fse03a3e	10/11 - p1 - casi d1.2.24 - had many fears
(74)	fgd04b1a	10/11 - casi d1.7.9 - problems feeling afraid or scared

ABS: Australian Bureau of Statistics; SEIFA: Socio-Economic Index for Areas.

Appendix C: Included variables and odds ratio estimates for Child Health CheckPoint response propensity models

Table C1: Adjusted odds ratios and 95% confidence intervals for LSAC wave 6 covariates associated with Child Health CheckPoint participation using backward stepwise variable selection with coefficient shrinkage.

Variable	Description	Backward Stepwise Maximum Likelihood shrinkage (0.91)
faanga	Parent 1 - angry parenting (v2)	1.121 (0.985, 1.276)
fairc	Parent 1 inductive reasoning scale (v3)	1.089 (0.989, 1.199)
ff11fm	Mother - language other than english spoken at home	0.809 (0.655, 0.999)
faparmonb	Parent 1 - parental monitoring version2	1.329 (1.088, 1.624)
fcnfsad2	Seifa - index of relative socio-econ advantage & disadvantage – 2011-SA2-score	1.003 (1.002, 1.004)
fhe09a4a	f2f m8.1.4 - extra curricular - art etc.	1.174 (1.017, 1.356)
fsc13a1zz	Parent 1 - f2f j14.3 - used for family in last 12 months	1.268 (1.090, 1.475)
fp2	Study child has 2 parents in the home	0.504 (0.347, 0.732)
ff03fp1	Parent 1 - age	1.029 (1.015, 1.043)
fnewpart	New partner of parent 1 in the home	0.612 (0.397, 0.944)
fhe09a	f2f m8.1- extra curricular - any	0.731 (0.544, 0.982)
fhe09a1a	f2f m8.1.1- extra curricular - community group	1.274 (1.045, 1.553)
ffd08a1	Parent 1 - school completion	0.865 (0.795, 0.941)
fpc65x9	Parent 1 - f2f k1.4.9 - after - school (last year) - parent lives/lived elsewhere	1.597 (1.021, 2.496)

Variable	Description	Backward Stepwise Maximum Likelihood shrinkage (0.91)
fabsra	Remoteness area (abs)	
	Major cities of Australia	1
	Inner regional Australia	0.734 (0.616, 0.874)
	Outer regional Australia	0.668 (0.523, 0.854)
	Remote Australia	0.200 (0.094, 0.428)
	Very remote Australia	0.198 (0.052, 0.745)
ffd08a2a	Parent 1 - comp other post-sec qualification	0.737 (0.610, 0.890)
fne09a3a	f2f m8.1.3- extra curricular - individual sport	0.811 (0.704, 0.935)
fstate	State of residence	
	NSW	1
	VIC	0.945 (0.785, 1.137)
	QLD	1.126 (0.929, 1.363)
	SA	1.773 (1.323, 2.376)
	WA	1.389 (1.084, 1.781)
	TAS	2.219 (1.450, 3.395)
	NT	5.627 (2.656, 11.923)
	ACT	0.950 (0.606, 1.488)
fbfath	Biological father in the home	1.584 (1.122, 2.235)
fpc65x3	Parent 1 - f2f k1.4.3 - after - school (last year) - program at study child's school	1.380 (1.097, 1.737)
fic08a3a	Parent 1 - overall school achievement	0.930 (0.859, 1.007)
fdisop1	Someone in home other than parent 1 has a disability	1.153 (0.986, 1.350)

ABS: Australian Bureau of Statistics; SEIFA: Socio-Economic Index for Areas.

Table C2: Adjusted odds ratios and 95% confidence intervals for LSAC wave 6 covariates associated with Child Health CheckPoint attendance at Main Assessment Centre using backward stepwise variable selection with coefficient shrinkage.

Variable	Description	Backward Stepwise Maximum Likelihood shrinkage (0.95)
fhe09a3a	f2f m8.1.3- extra curricular - individual sport	0.826 (0.713, 0.957)
fairc	Parent 1 - parent 1 inductive reasoning scale (v3)	1.115 (1.006, 1.237)
fhe09a4a	f2f m8.1.4 - extra curricular - art etc.	1.352 (1.164, 1.571)
faparmonb	Parent 1 - parental monitoring version2	1.508 (1.221, 1.861)
fcnfsad2	Seifa - index of relative socio-econ advantage & disadvantage – 2011-SA2-score	1.002 (1.001, 1.003)
fstate	State of residence	
	NSW	1
	VIC	0.903 (0.744, 1.096)
	QLD	0.794 (0.647, 0.975)
	SA	1.552 (1.151, 2.093)
	WA	1.187 (0.921, 1.531)
	TAS	0.026 (0.004, 0.191)
	ACT	1.337 (0.851, 2.103)
fpc65x3	Parent 1 - f2f k1.4.3 - after - school (last year) - program at study child's school	1.307 (1.038, 1.647)
ffd08a1	Parent 1 - school completion	0.828 (0.753, 0.911)
ff03fp1	Parent 1 - age	1.029 (1.014, 1.044)
ffd08a2a	Parent 1 - comp other post-sec qualification	0.704 (0.569, 0.871)
flc08a3a	Parent 1 - overall school achievement	0.925 (0.851, 1.006)

Variable	Description	Backward Stepwise Maximum Likelihood shrinkage (0.95)
fhe09a3a	f2f m8.1.3- extra curricular - individual sport	0.826 (0.713, 0.957)
fhe09a1a	f2f m8.1.1- extra curricular - community group	1.527 (1.244, 1.874)
fsc12a1w	Parent 1 - f2f j14.1.5- used for sc - disability services	0.460 (0.209, 1.016)
fpc65x9	Parent 1 - f2f k1.4.9 - after - school (last year) - parent lives/lived elsewhere	1.533 (0.986, 2.383)
fabsra	Remoteness area (abs)	
	Major cities of Australia	1
	Inner regional Australia	0.556 (0.459, 0.672)
	Outer regional Australia	0.277 (0.200, 0.383)
	Remote Australia	0.115 (0.034, 0.390)
	Very remote Australia	0.343 (0.091, 1.288)
fsc13a1zz	Parent 1 - f2f j14.3 - used for family in last 12 months	1.188 (1.015, 1.390)
ff11fp1	Parent 1 - language other than english spoken at home	0.661 (0.525, 0.833)
fbmoth	Biological mother in the home	0.440 (0.211, 0.917)
fdisop1	Someone in home other than parent 1 has a disability	1.162 (0.983, 1.373)

ABS: Australian Bureau of Statistics; SEIFA: Socio-Economic Index for Areas.

Table C3: Adjusted odds ratios and 95% confidence intervals for LSAC wave 6 covariates associated with Child Health CheckPoint attendance at Main Assessment Centre/Mini Assessment Centre using backward stepwise variable selection with coefficient shrinkage.

Variable	Description	Backward Stepwise Maximum Likelihood shrinkage (0.94)
ff11fm	Mother - language other than english spoken at home	0.658 (0.529, 0.819)
fhe09a4a	f2f m8.1.4 - extra curricular - art etc.	1.274 (1.106, 1.468)
fsc12a1w	Parent 1- f2f j14.1.5- used for sc - disability services	0.517 (0.251, 1.064)
faparmonb	Parent 1 - parental monitoring version2	1.376 (1.126, 1.682)
fcnfsad2	Seifa - index of relative socio-econ advantage & disadvantage – 2011-SA2-score	1.002 (1.001, 1.003)
fhe09a3a	f2f m8.1.3- extra curricular - individual sport	0.773 (0.671,0.889)
fsc12a1z	Parent 1- f2f j14.1 - used services for sc in last year	0.749 (0.582, 0.965)
ffd08a1	Parent 1 - school completion	0.832 (0.76
ffd08a1	Parent 1 - school completion	0.832 (0.762, 0.910)
ff03fp1	Parent 1 - age	1.027 (1.013, 1.042)
fsc13a1zz	Parent 1 - f2f j14.3 - used for family in last 12 months	1.209 (1.042, 1.403)
zf12fp1	Parent 1 - indigenous status	0.443 (0.217, 0.906)
fhe09a1a	f2f m8.1.1- extra curricular - community group	1.481 (1.216, 1.802)
fstate	State of residence	
	NSW	1
	VIC	0.983 (0.813, 1.189)
	QLD	1.074 (0.882, 1.308)
	SA	1.567 (1.169, 2.101)

Variable	Description	Backward Stepwise Maximum Likelihood shrinkage (0.94)
	WA	1.495 (1.165, 1.919)
	TAS	2.070 (1.343, 3.188)
	NT	3.245 (1.560, 6.754)
	ACT	1.530 (0.978, 2.392)
fpc09	Care at any time - total hours per week	1.020 (1.005, 1.035)
ffd08a2a	Parent 1 - comp other post-sec qualification	0.644 (0.527, 0.787)
fabsra	Remoteness area (abs)	
	Major cities of Australia	1
	Inner regional Australia	0.686 (0.572, 0.823)
	Outer regional Australia	0.539 (0.414, 0.701)
	Remote Australia	0.085 (0.029, 0.252)
	Very remote Australia	0.304 (0.079, 1.167)

ABS: Australian Bureau of Statistics; SEIFA: Socio-Economic Index for Areas.

Table C4: Adjusted odds ratios and 95% confidence intervals for LSAC wave 6 covariates associated with Child Health CheckPoint study child blood sample using backward stepwise variable selection with coefficient shrinkage.

Variable	Description	Backward Stepwise Maximum Likelihood shrinkage (0.92)
fhe09a1a	f2f m8.1.1- extra curricular - community group	1.410 (1.153, 1.725)
fairc	Parent 1 - parent 1 inductive reasoning scale (v3)	1.123 (1.012, 1.247)
fsc13a1zz	Parent 1 - f2f j14.3 - used for family in last 12 months	1.186 (1.012, 1.389)
faparmonb	Parent 1 - parental monitoring version2	1.398 (1.132, 1.727)
fcnfsad2	Seifa - index of relative socio-econ advantage & disadvantage - 20	1.002 (1.001, 1.003)
fabsra	Remoteness area (abs)	
	Major cities of Australia	1
	Inner regional Australia	0.813 (0.674, 0.982)
	Outer regional Australia	0.429 (0.320, 0.577)
	Remote Australia	0.100 (0.033, 0.305)
	Very remote Australia	0.396 (0.104, 1.513)
fsc12a1w	Parent 1 - f2f j14.1.5- used for sc - disability services	0.223 (0.078, 0.640)
fhe09a3a	f2f m8.1.3- extra curricular - individual sport	0.800 (0.690, 0.926)
ff03fp1	Parent 1 - age	1.030 (1.015, 1.046)
fstate	State of residence	
	NSW	1
	VIC	0.895 (0.733, 1.092)
	QLD	1.009 (0.822, 1.238)
	SA	1.506 (1.115, 2.035)
	WA	1.240 (0.959, 1.603)

Variable	Description	Backward Stepwise Maximum Likelihood shrinkage (0.92)
	TAS	2.052 (1.314, 3.205)
	NT	4.373 (2.089, 9.151)
	ACT	1.118 (0.711, 1.759)
fhe09a4a	f2f m8.1.4 - extra curricular - art etc.	1.246 (1.075, 1.445)
fnsib	Number of siblings of study child in household	1.078 (0.999, 1.163)
zf12fp1	Parent 1 - indigenous status	0.440 (0.202, 0.959)
fpc09	Care at any time - total hours per week	1.022 (1.007, 1.038)
ffd08a2a	Parent 1 - comp other post-sec qual	0.695 (0.562, 0.859)
fsc12a1z	Parent 1 - used services for sc in last year	0.780 (0.597, 1.019)
fp2	Study child has 2 parents in the home	0.772 (0.619, 0.962)
fsatir	Sati reactivity	0.888 (0.805, 0.980)
ffd08a1	Parent 1 - school completion	0.852 (0.774, 0.938)
ff11fm	Mother - language other than english spoken at home	0.556 (0.437, 0.707)

ABS: Australian Bureau of Statistics; SEIFA: Socio-Economic Index for Areas.

Table C5: Adjusted odds ratios and 95% confidence intervals for LSAC wave 6 covariates associated with Child Health CheckPoint attending parent blood sample using backward stepwise variable selection with coefficient shrinkage.

Variable	Description	Backward Stepwise Maximum Likelihood shrinkage (0.91)
fhe09a1a	f2f m8.1.1- extra curricular - community group	1.410 (1.156, 1.719)
fairc	Parent 1- parent 1 inductive reasoning scale (v3)	1.124 (1.016, 1.244)
fpc65x9	Parent 1 - f2f k1.4.9 - after - school (last year) - parent lives/lived elsewhere	1.541 (0.992, 2.394)
faparmonb	Parent 1 - parental monitoring version2	1.406 (1.144, 1.727)
fcnfsad2	Seifa - index of relative socio-economic advantage and disadvantage - 2011-SA2-score	1.002 (1.001, 1.003)
fhe09a2a	f2f m8.1.2 - extra curricular - team sport	0.874 (0.750, 1.018)
fhe09a4a	f2f m8.1.4 - extra curricular - art etc.	1.162 (1.000, 1.348)
ffd08a1	Parent 1 - school completion	0.846 (0.772, 0.929)
ff03fp1	Parent 1 - age	1.033 (1.018, 1.048)
fsc13a1zz	Parent 1 - f2f j14.3 - used for family in last 12 months	1.210 (1.038, 1.410)
fabsra	Remoteness area (abs)	
	Major cities of Australia	1
	Inner regional Australia	0.608 (0.466, 0.793)
	Outer regional Australia	0.287 (0.187, 0.442)
	Remote Australia	0.027 (0.007, 0.109)
	Very remote Australia	0.133 (0.024, 0.743)
fhe09a3a	Extra curricular - individual sport	0.819 (0.709, 0.945)
ff11fm	Mother - language other than english spoken at home	0.573 (0.456, 0.721)
fpc09	Care at any time - total hours per week	1.017 (1.002, 1.033)

Variable	Description	Backward Stepwise Maximum Likelihood shrinkage (0.91)
fremote	Remoteness area classification (aria)	
	Highly accessible	1
	Accessible	0.964 (0.776, 1.197)
	Moderately Accessible	1.450 (0.997, 2.109)
	Remote	1.864 (0.925, 3.756)
	Very remote	2.375 (0.807, 6.984)
	Not determined	1.985 (0.835, 4.720)
fp2	Study child has 2 parents in the home	0.758 (0.608, 0.945)
fstate	State of residence	
	NSW	1
	VIC	1.033 (0.850, 1.126)
	QLD	1.034 (0.844, 1.268)
	SA	1.691 (1.252, 2.282)
	WA	1.357 (1.051, 1.752)
	TAS	2.706 (1.746, 4.193)
	NT	4.679 (2.215, 9.884)
	ACT	1.374 (0.879, 2.147)
zf02m1	Study child - sex	1.137 (0.981, 1.317)
ffd08a2a	Parent 1 - completed other post-secondary qualification	0.706 (0.574, 0.867)
zf12fp1	Parent 1 - indigenous status	0.425 (0.202, 0.894)
ffd09a1a	Parent 1- currently studying	1.234 (0.999, 1.520)

ABS: Australian Bureau of Statistics; SEIFA: Socio-Economic Index for Areas.

Appendix D: Distribution of the five final sets of survey weights

Figure D1: Distribution of final survey weights for Child Health CheckPoint overall participation.

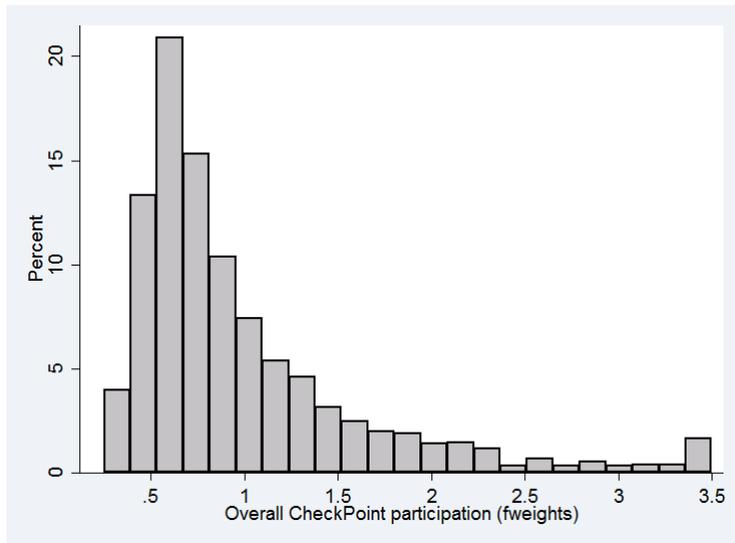


Figure D2: Distribution of final survey weights for Child Health CheckPoint attendance at a Main Assessment Centre participation.

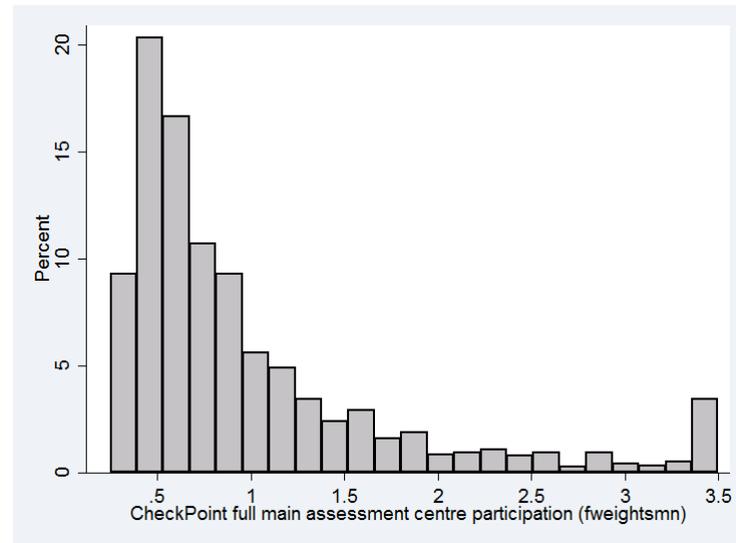


Table D1: Summary of Checkpoint overall participation weights (fweightscp)

Mean	Standard Deviation	Minimum	Maximum	Range	Sum	Count
1.00	0.6459	0.25	3.5	3.25	1874	1874

Table D2: Summary of Checkpoint Main Assessment Centre participation weights (fweightsmn)

Mean	Standard Deviation	Minimum	Maximum	Range	Sum	Count
1.00	0.7611	0.25	3.5	3.25	1356	1356

Figure D3: Distribution of final survey weight for Child Health CheckPoint attendance at a Main Assessment Centre or Mini Assessment Centre participation.

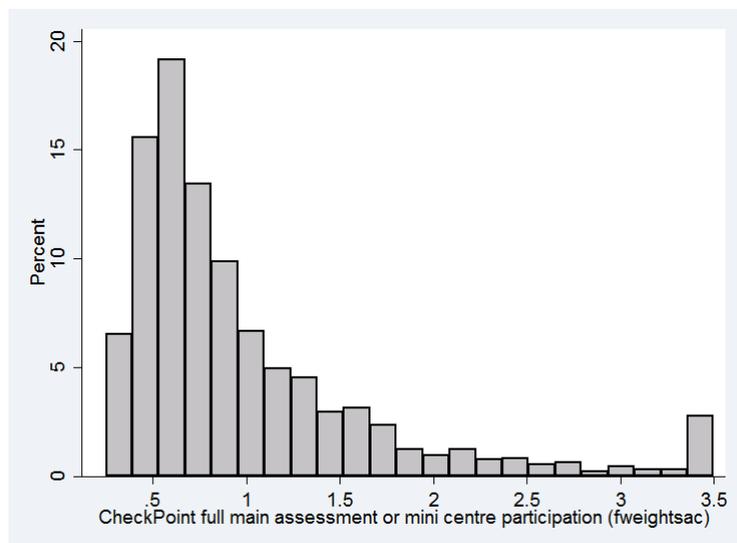


Table D3: Summary of Checkpoint Main Assessment Centre or Mini Assessment Centre participation weight (fweightsac)

Mean	Standard Deviation	Minimum	Maximum	Range	Sum	Count
1.00	0.6986	0.25	3.5	3.25	1509	1509

Figure D4: Distribution of final survey weight for Child Health CheckPoint study child blood sample participation.

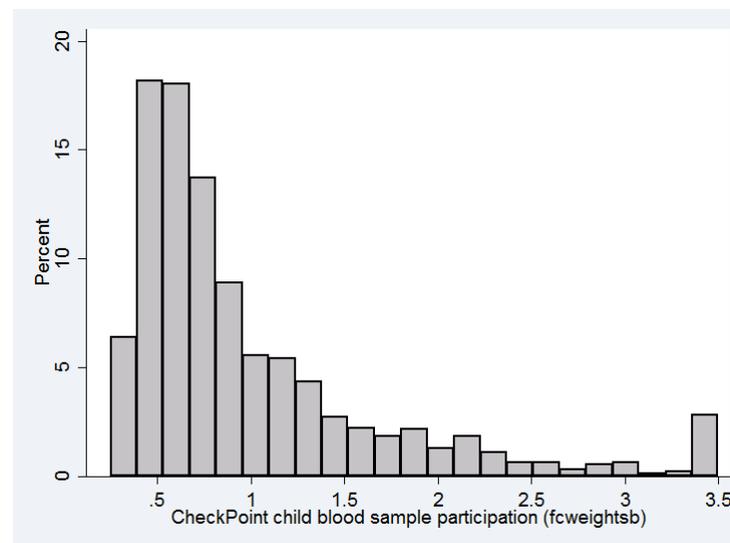


Table D4: Summary of Checkpoint study child blood sample participation weights (fweightsb)

Mean	Standard Deviation	Minimum	Maximum	Range	Sum	Count
1.00	0.7117	0.25	3.5	3.25	1237	1237

Figure D5: Distribution of final survey weight for Child Health CheckPoint attending parent blood sample participation.

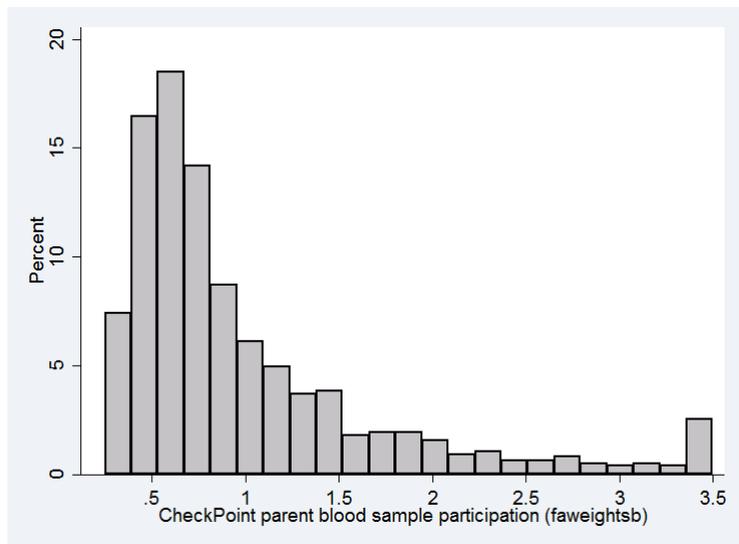


Table D5: Summary of Checkpoint attending parent blood sample participation weight (faweightsb)

Mean	Standard Deviation	Minimum	Maximum	Range	Sum	Count
1.00	0.7147	0.25	3.5	3.25	1373	1373