



Growing Up in Scotland Sweep 12: 2025

User Guide

Eleanor Holman, Helena Wilson, Paul Bradshaw and Anna Keyes, Scottish
Centre for Social Research

Scottish Centre for Social Research
Scotiabank House
6 South Charlotte Street
Edinburgh, EH2 4AW
T 0131 240 0210
www.scotcen.org.uk

A Company Limited by Guarantee. Registered in England No.4392418.
A Charity registered in England and Wales (1091768) and Scotland
(SC038454)

Contents

1	Survey details	4
1.1	Study aims and objectives	4
1.2	Study design	4
1.3	Sample design	5
1.3.1	BC1 Main sample (the original birth cohort)	5
1.3.2	BC1 Boost sample (refreshment sample recruited in 2018)	6
1.4	Sweep 12 data collection elements	7
1.5	Developing and piloting	7
1.6	Sweep 12 fieldwork timing	8
1.7	Incentives and thank you gifts	8
1.8	Response	8
1.9	Length of interview	10
2	The interview	11
3	Coding and editing	12
4	Weighting the data	13
4.1	Background	13
4.2	Weights developed for sweep 12	13
4.2.1	Sample definitions	13
4.2.2	Longitudinal weights	14
4.2.3	Cross-sectional weights	15
4.3	Applying the weights	15
5	Using the data	17
5.1	Variables on the data file	17
5.2	Variable naming convention	17
5.3	Variable labels	18
5.4	Derived variables	18
5.5	Multicoded questions	18
5.6	Indicators and summary variables	19
5.6.1	Household details collected at sweep 12	19
5.6.2	Household data	19
5.6.3	National Statistics Socio-economic Classification (NS-SEC)	19
5.6.4	Area-level variables	20
5.6.5	Strengths and Difficulties Questionnaire (SDQ)	22
5.6.6	Mental wellbeing: selected items from the Students' Life Satisfaction Scale	22
5.6.7	Anxiety: Generalised Anxiety Disorder Assessment (GAD-2)	23
5.6.8	Depression: Selected items from the Patient Health Questionnaire (PHQ-9)	23

5.6.9	Alcohol and smoking.....	24
5.6.10	Drug use	25
5.6.11	Anti-social behaviour and offending	25
5.6.12	Parent-Child Attachment: selected items from the People In My Life (PIML) scale.....	26
5.6.13	Peer Attachment: selected items from the People In My Life (PIML) scale	26
5.6.14	Additional new measures sourced elsewhere	27
5.7	Dropped variables.....	27
5.8	Missing values conventions	28
5.9	Specific issues to note	28
6	Documentation	29
7	Contact details	30

1 Survey details

1.1 Study aims and objectives

The overarching aim of the Growing Up in Scotland study is set out in its purpose, which is:

“To generate, through robust methods, specifically Scottish data about outcomes throughout childhood and into adulthood for children growing up in Scotland across a range of key domains:

- *Cognitive, social, emotional and behavioural development*
- *Physical and mental health and wellbeing*
- *Childcare, education and employment*
- *Home, family, community and social networks*
- *Involvement in offending and risky behaviour*

Such data will encompass, in particular, topics where Scottish evidence is lacking and policy areas where Scotland differs from the rest of the UK.”

1.2 Study design

GUS was initially based on two cohorts of children: the first aged approximately 10 months at the time of first interview (involving around 5217 children at the first sweep) and the second aged approximately 34 months (involving around 2800 children at the first sweep). In 2018, an additional 502 families were recruited to the original birth cohort. Further details are provided in section 1.3.2.

A second birth cohort of 6127 children aged around 10 months at the first interview was recruited in 2011.

The configuration of cohorts and sweeps for all sweeps of data collection launched to date is summarised in Table 1.1. BC1 refers to the younger of the two original cohorts ('birth cohort 1'), CC to the slightly older cohort ('child cohort') and BC2 to the later birth cohort ('birth cohort 2').

A key aim of using multiple cohorts was to allow the study to provide three types of data:

- Cross-sectional time specific data – e.g. what proportion of 17-18 year-olds were living in single parent families in 2021-23?
- Cross-sectional time series data – e.g. is there any change in the proportion of 5-year-old children living in single parent families between 2010 and 2016?
- Longitudinal cohort data – e.g. what proportion of children who were living in single parent households aged 10-11 are living in different family circumstances at the time they are aged 17-18?

Table 1.1 Ages, stages and timing of data collection sweeps by cohort				
Child's age* (school stage)	Cohort, year of data collection (sweep) <i>achieved sample size</i>			
	Child cohort	Birth cohort 1		Birth Cohort 2
		Main sample	Boost sample	
10 months	-	2005/06 (1) 5217	-	2011 (1) 6127
1-2 years	-	2006/07 (2) 4512	-	-
2-3 years	2005/06 (1) 2858	2007/08 (3) 4193	-	2013 (2) 5020
3-4 years	2006/07 (2) 2500	2008/09 (4) 3994	-	2014** (2.5) 3237
4-5 years	2007/08 (3) 2332	2009/10 (5) 3833	-	2015 (3) 4434
5-6 years	2008/09 (4) 2200	2010/11 (6) 3657	-	-
7-8 years	-	2012/13 (7) 3453	-	-
8-9 years (Primary 5)	-	2013/14** (7.5) 2775	-	-
9-10 years (Primary 6)	-	2014/15 (8) 3151	-	-
10-11 years (Primary 7)	-	2015/16** (8.5) 2099	-	-
12-13 years (Secondary 1)	-	2017/18 (9) 2917	502	-
14-15 years (Secondary 3)	-	2019/20 (10) 2669	274	-
17-18 years (Secondary 6)	-	2021-23 (11) 2384	265	-
19-20 years	-	2025 (12) 2005	201	-

* Up to age 8, data collection was timed for each case so that data was captured from children when they were of the same chronological age. Typically the spread was around 3 months. From age 8 onwards, the timing aimed to capture children at the same school stage introducing a wider age range at the point of data collection.

**This data was collected via a web questionnaire or telephone interview.

1.3 Sample design

1.3.1 BC1 Main sample (the original birth cohort)

The original or 'main' BC1 sample was recruited at sweep 1.

The initial area-level sampling frame was created by aggregating Data Zones. Data Zones are small geographical output areas created for the Scottish Government. Data Zones are used by Scottish Neighbourhood Statistics to release small area statistics. The Data Zone geography covers the whole of Scotland. The geography is hierarchical, with Data Zones nested within Local Authority boundaries. Each data zone contains between 500 and 1,000 household residents. More information can be found on the Scottish Government's Small Area Statistics website: [Small area statistics: Data Zones and Intermediate Zones - gov.scot](https://smallareastatistics.gov.scot)

The Data Zones were aggregated to give an average of 57 births per area per year (based on the average number of births in each Data Zone for the preceding 3 years). It was estimated that this number per area would provide us with the required sample size. Once the merging task was complete, the list of aggregated areas was sorted by Local Authority¹ and then by the Scottish Index of Multiple Deprivation Score (SIMD). 130 areas were then selected at random. The Department of Work and Pensions then sampled children from these 130 sample points.

Within each sample point, the Child Benefit records were used to identify all babies and three-fifths of toddlers who were born between 1st June 2004 and 31st May 2005. The sampling of children was carried out on a month-by-month basis in order to ensure that the sample was as complete and accurate as possible at time of interview.

In cases where there was more than one eligible child in the selected household, one child was selected at random. If the children were twins, they had an equal chance of being selected. If the eligible children were in different age cohorts the younger child had a higher chance of being selected given that those children had a higher chance of being included in the sample overall.

After selecting the eligible children, the DWP made a number of exclusions before transferring the sample details. These exclusions included cases they considered 'sensitive' and children that had been sampled for research by the DWP in the last 3 years.

1.3.2 BC1 Boost sample (refreshment sample recruited in 2018)

Whilst the overall levels and patterns of attrition seen in GUS are typical for a cohort study of its kind, the effects of attrition are spread unevenly over the sample, with some sub-groups affected more than others. Analysis of the achieved sample from Birth Cohort 1 after sweep 8 revealed that two groups in particular had become under-represented since the beginning of the study: children born to mothers aged 16-24 at time of birth and children living in the 15% most deprived areas (according to the Scottish Index of Multiple Deprivation).

To resolve this under-representation, a boost sample for BC1 was recruited to the study as part of phase 2 fieldwork for sweep 9, specifically targeting families in the under-represented groups. Like the sample for the original birth cohort, the sample for the boost was drawn from Child Benefit records held by HMRC. For further details about the sampling approach and rationale, please see the GUS BC1 sweep 9 User Guide ([Sweep 9 User Guide \(PDF\)](#))

At sweep 9, a total of 502 families were recruited to the boost sample. All families who consented to follow-up and who had not subsequently withdrawn from the study were issued for sweep 12 fieldwork (n=430).

¹ Local Authority has been used as a stratification variable during sampling, this means the distribution of the GUS sample by Local Authority will be representative of the distribution of Local Authorities in Scotland. However, the sample sizes are such that we would not recommend analysis by Local Authority. The small sample sizes would give misleading results.

1.4 Sweep 12 data collection elements

The Scottish Centre for Social Research (ScotCen) was commissioned by the Scottish Government to undertake the design and delivery of sweep 12. At this sweep, data was only collected from the cohort member (young person). There was no parent questionnaire/interview or other measurements taken. The data was collected either **online** or **in-person** using the **same questionnaire**. Cohort members were initially asked to complete the survey online. Those who did not were issued for a face-to-face interview.

For those who completed the survey with an interviewer, there was both a Computer Assisted Personal Interview (CAPI) with questions asked by the interviewer and a self-complete instrument (Computer Assisted Self Interview – CASI) for the cohort member to complete on the interviewer's laptop during the visit. For those who completed the survey online, these two elements were delivered as a web survey (Computer Assisted Web Interview – CAWI).

1.5 Developing and piloting

Policy priorities and key topics of interest for the sweep 12 questionnaire were initially discussed and agreed with the study's Scottish Government Project Manager and a number of internal and external stakeholder groups, including the GUS Questionnaire Advisory Group.

Informal qualitative consultation was also undertaken with groups of young people who were the same age cohort members would be at the time of data collection. This took the form of small, in-person discussion groups facilitated by the ScotCen research team. The aim was to find out what the topical issues were for young people. A total of three groups were convened: one with university students, one with college students, and one with young people who were in employment, training, or looking for work. Feedback from stakeholders and young people was then used by the GUS team at ScotCen to develop the questionnaires with input from the study's Questionnaire Advisory Group and policy teams across the Scottish Government.

Cognitive testing of selected new items in the young person questionnaire was carried out in August 2024. Cognitive interviews were undertaken with 11 young people of a similar age to cohort members and with a range of characteristics representing key expected differences in the sample at this sweep. In particular, the cognitive pilot sample included some young people who were employed full-time, and some who were in further or higher education. A full dress rehearsal pilot, including near-to-final versions of the cohort member CAWI, CASI and CAPI questionnaires, was conducted in October 2024. The existing GUS pilot sample was used and a total of 82 cases were issued. Overall, 18 young person CAPI interviews were conducted (out of 60 cases issued to field, 30% response rate) and 27 young person CAWI questionnaires were completed (either before or during the CAPI phase, 33% response).

All piloted instruments and procedures, and the final research protocols were reviewed and approved by the National Centre for Social Research Research Ethics Committee.

1.6 Sweep 12 fieldwork timing

In sweeps 1-7, fieldwork was conducted over a 14-month period with cases issued to field according to the child's age and interviews taking place around a specified date calculated according to the child's birthday (the 'target interview date'). Ahead of sweep 8 there was interest in interviewing families according to the child's school year. Therefore, from sweep 8, fieldwork moved from an 'ages' to a 'stages' approach. This meant that the age gap between the young people at the time of interview was larger at sweeps 8-11 than at previous sweeps.

Because of how children were initially sampled, cohort members in BC1 span two different school years. In sweeps 8-10 the fieldwork was conducted over two phases coinciding with when the cohort would be in their first or second term of their school year. Fieldwork in sweep 11 was delayed because of covid restrictions, so Phase 1 fieldwork took place between November 2021 and November 2022 and Phase 2 fieldwork took place between September 2022 and June 2023.

In sweep 12, fieldwork was also in two phases, though the phases were differentiated by mode of data collection rather than school stages². In the first phase, cohort members were invited to take part online and in the second phase, those who hadn't taken part online were issued to face-to-face fieldwork. The dates were:

- Letters (with leaflet), emails and text messages sent to cohort members in February 2025, inviting them to take part online.
- Letter, email and text reminders were sent in March 2025 to cohort members who had not completed their web questionnaire.
- Face-to-face fieldwork took place between May and October 2025. The online questionnaire remained open during this phase so participants could still complete the survey online if they chose.

1.7 Incentives and thank you gifts

For the web phase, participants in the boost sample were offered a £20 Love2Shop voucher, as they were disproportionately from disadvantaged backgrounds at higher risk of non-response. This mirrors the incentive approach which has applied to the boost sample since they were introduced to the study in sweep 9.

In the face-to-face fieldwork phase, all cohort members were offered a £20 Love2Shop e-voucher if they completed an interview – including if they completed a questionnaire online during the face-to-face phase.

All cohort members who participated at sweep 12 were sent a thank you email or letter.

1.8 Response

² Having all left school by sweep 12, there was no continuing need to align the timing of data collection with a cohort member's school stage.

Table 1.2 shows historical response/attrition for BC1. At sweep 12, a total of 2005 interviews were achieved with the original sample through either web or face-to-face data collection. This represents 38% of cases achieved at sweep 1.

Table 1.2 BC1 historical response		
	Cases achieved	% of sweep 1 cases
Sweep 1	5217	-
Sweep 2	4512	86%
Sweep 3	4193	80%
Sweep 4	3994	77%
Sweep 5	3833	73%
Sweep 6	3657	70%
Sweep 7	3456	66%
Sweep 8	3150	60%
Sweep 9 (Main sample only)	2917	56%
Sweep 10 (Main sample only)	2669	51%
Sweep 11 (Main sample only)	2384	46%
Sweep 12 (Main sample only)	2005	38%

Details of the number of cases issued and achieved at sweep 12 for both samples are presented in Table 1.3.

Table 1.3 Sweep 12 response by sample type						
	Total cross-sectional sample (Main and Boost samples)		Main sample		Boost sample	
	No. of cases	% of issued in-scope	No. of cases	% of issued in-scope	No. of cases	% of issued in-scope
Total in-scope/issued	3373	100%	2943	100%	430	100%
Total achieved	2206	65%	2005	68%	201	47%
<i>Online</i>	1268	33%	1157	37%	111	6%
<i>Face-to-face</i>	938	32%	848	31%	90	42%
Total with complete data	2163	64%	1964	67%	199	46%
<i>Online</i>	1225	36%	1116	38%	109	25%
<i>Face-to-face</i>	938	28%	848	29%	90	21%
Total with partial data	43	1%	41	1%	2	>1%
<i>Online</i>	43	1%	41	1%	2	>1%
<i>Face-to-face</i>	0	0%	0	0%	0	0%

The variable DMode can be used to determine which cases responded by CAWI or CAPI.

Table 1.4 provides further details on the reasons for non-response amongst those cases which did not result in a productive interview with the cohort member in either the web or face-to-face phases.

Table 1.4 Sweep 12 CAPI non-response by sample type

	Total cross-sectional sample (Main and Boost samples)		Main sample		Boost sample	
	No. of cases	% of issued in-scope	No. of cases	% of issued in-scope	No. of cases	% of issued in-scope
Total issued to CAPI	2439	100%	2101	100%	338	100%
CAPI interview achieved	938	38%	848	40%	90	27%
CAWI interview in CAPI phase	355	15%	331	16%	24	7%
Total unproductive	1146	47%	922	44%	224	66%
<i>Non-contact*</i>	295	12%	244	12%	51	15%
<i>Movers**</i>	235	10%	183	9%	52	15%
<i>Refusal</i>	536	22%	428	20%	108	32%
<i>Other</i>	80	3%	67	3%	13	4%

*In non-contact cases, the interviewer was unable to establish contact with anyone at the issued address.

**Movers were cases where the interviewer established contact at the issued address but determined the cohort member was no longer living there and was otherwise unable to make contact with them.

As the table shows, the vast majority of unproductive cases were as a result of a refusal by the cohort member (or their proxy³, usually a parent/carer) to take part. There were also a large number of cases where the cohort member had moved and could not be traced. This reflects the particular age and stage of the cohort at this sweep where many young people were moving out of their parental home, including to attend university or to work or study abroad.

1.9 Length of interview

The median interview length was 32 minutes for the online questionnaire and 43 minutes for the face-to-face questionnaire.

³ In some circumstances, when attempting to make contact, interviewers were not always able to speak to the cohort member but instead spoke to another member of the household. In some such cases, this person – a 'proxy' for the cohort member – refused on the cohort member's behalf.

2 The interview

In sweep 12, only the cohort member was interviewed. This was the sixth time they were directly interviewed. For details about consent procedures and contact procedures please see the interviewer instructions. Copies of the invitation letters and other documents are provided in the data documentation.

Questionnaires were either completed online or face-to-face, with the same questions asked in both. For the face-to-face interview, there was an interviewer led section (CAPI) and a self-completion section (CASI).

Participants were first invited to take part online with an advance letter, email and text. During online fieldwork, participants were also sent two reminders via email and text and one reminder letter.

After the online fieldwork had closed, all participants who hadn't completed the online questionnaire or withdrawn from the study were issued to face-to-face fieldwork. These participants were sent an email telling them an interviewer would be visiting them to complete their questionnaire in person. During the face-to-face fieldwork participants were also sent two further reminders by email and text and a final reminder letter.

During the face-to-face fieldwork period, the online questionnaire was reopened so participants could chose to do it online rather than in person if they preferred. If a case completed online, interviewers were told and cases were removed from face-to-face fieldwork, but there were some cases where both an online and face-to-face interview was completed. For these cases, because the face-to-face data was more complete (i.e. fewer variables with missing data) and more current (with face-to-face interviews taking place after the web phase), the face-to-face interview has been included in the dataset.

3 Coding and editing

Additional coding and editing tasks were performed after the interviews were conducted. The accompanying **CAPi edit instructions** provide details of the tasks that were conducted.

Coding and editing tasks were carried out on data collected across both modes.

4 Weighting the data

4.1 Background

As noted above, only the cohort young person was invited to participate at sweep 12. Therefore, unlike previous sweeps, weights for main carer responses were not required. The methods used for the young person weights were nonetheless kept consistent with previous sweeps.

4.2 Weights developed for sweep 12

Two sets of weights were generated for the sweep 12 data:

- A longitudinal weight for responding cohort young people who had also responded to sweep 11 and whose main carer had responded at every other sweep of GUS up to and including sweep 10.
- A cross-sectional weight for all cohort young people responding at sweep 12.

Data collection from cohort young people was undertaken using both web and face-to-face interviews. For the purposes of weighting, responses were treated in exactly the same way regardless of the interview mode.

4.2.1 Sample definitions

The sweep 12 sample of young person respondents can be split into three groups. For the purposes of describing the weighting, these have been named Sample A, Sample B, and Boost Sample. They are defined as follows:

- Sample A are young people whose main carers responded to all previous sweeps up to sweep 10 and who took part in sweeps 11 and 12 themselves.
- Sample B are young people who took part in sweep 12 themselves and whose main carers responded to sweep 1 but whose carers did not participate in one or more of sweeps 2 to 11.
- Boost sample are young people who were part of the refreshment sample added at sweep 9 and took part in sweep 12 themselves.

The three samples were treated separately during the weighting, for consistency with the weighting of previous sweeps.

- There were 1830 young people in Sample A, 1476 of whom responded at sweep 12. The response rate was 81%.
- There were 1126 young people in Sample B, 529 of whom responded at sweep 12. The response rate was 47%.
- There were 430 young people in the boost sample, 201 of whom responded at sweep 12. The response rate was 47%.

The issued and responding sample sizes for the three groups in the young person data are given in Table 4.1 below.

Table 4.1 Young person sample groups for weighting and corresponding response figures

Sample	No. issued	No. responding	Response rate
Sample A	1817	1476	81%
Sample B	1126	529	47%
Boost	430	201	47%
Combined (A+B+Boost)	3373	2206	65%

Only members of Sample A, whose main carers responded to all previous GUS sweeps, received a longitudinal weight. This weight is described in greater detail in section 4.2.2 below.

All young people that responded in sweep 12 were given a cross-sectional weight, including Sample A, Sample B, and Boost Sample cases. This weight is described in greater detail in section 4.2.3 below.

4.2.2 Longitudinal weights

Young person longitudinal weights were only generated for responding young people in Sample A. The weights were generated using a model-based method with data from previous sweeps. This same method was previously used for the main carer longitudinal weights. Ineligible households were not included in the modelling. The base for the model was the sweep 11 young person longitudinal respondents ($n = 1830$), including 13 cases that were not issued in sweep 12 as they withdrew from the study.

A model of whether the young person responded to sweep 12 was fitted using forward and backward stepwise logistic regression. The potential predictors tested were a set of socio-demographic variables collected from previous sweeps of the study, including both characteristics of main carers and of the household. The non-response model was run weighted by the sweep 11 young person longitudinal sample weights.

The following variables were significant in the stepwise regressions and therefore included in the final model: area deprivation (SIMD quintile), household tenure, household income, main carer National Statistics Socio-Economic Classification (NS-SEC) category, main carer employment status, household employment status, and mother's employment status.

The non-response weight was calculated as the inverse of the modelled probabilities of response. This was trimmed at the 99th percentile to improve efficiency and multiplied by the sweep 11 young person longitudinal sample weight. The combined weight was checked for outliers and the top 3 weights trimmed. The final weights were scaled to the sweep 12 young person Sample A size of 1476. They have an efficiency of 68% and an effective sample size of 1008.

4.2.3 Cross-sectional weights

Cross-sectional weights were generated for all young people who responded to sweep 12. The cross-sectional responding sample included Sample A, Sample B, and Boost Sample cases.

Calibration weighting was applied to the young person combined sample to create the cross-sectional weights. This method adjusts a set of start weights using an iterative procedure so that they match pre-defined population totals. The resulting weights, when applied to the young person data, produce survey estimates that match the population estimates.

Population estimates for calibrating the combined young person responding sample were derived from weighting young person Sample A by the young person longitudinal weight. Variables included in the calibration were selected by fitting a model using forward and backward stepwise regression. The base of the model was the young person cross-sectional responding sample and the outcome variable was response to all previous waves. In other words, whether the young person was interviewed in sweep 11 and their main carer also interviewed in sweeps 1 to 10. The models were run weighted by a start weight, consisting of the young person longitudinal weight for Sample A cases and the cross-sectional weights from the wave in which the young person (sweep 11) or main carer (earlier waves) last responded for Sample B and Boost cases.

The potential predictors tested were a set of socio-demographic variables collected in sweep 12 of the study, including both characteristics of the young person and of the area in which they live. These measures covered a consistent set of socio-demographic characteristics to the main carer measures used in cross-sectional weighting of previous waves.

The following variables were significant in the stepwise regressions and therefore included in the calibration: whether the young person received income from their parents; whether the young person received income from student loans; the young person's general health; whether the young person was able to save on a regular basis; quintiles of SIMD education domain scores, and key subgroup⁴. Three additional key variables were included in the calibration - despite not being found significant during modelling - because of their importance for analysis, bias reduction, and for continuity with previous sweeps: household tenure, urban-rural status (of home address), and area deprivation (SIMD quintile).

After calibration, the weights were checked for outliers and the top four weights trimmed. The final weights were scaled to the sweep 12 young person responding sample size of 2206. They have an efficiency of 74% and an effective sample size of 1622.

4.3 Applying the weights

The cross-sectional weights should be used for any cross-sectional analysis of all sweep 12 young person data. All respondents that responded at sweep 12,

⁴ A four-category variable created when introducing the boost sample to the study identifying whether the cohort child's parent was under 25 when they were born and whether their address was in the most deprived 15% of areas in Scotland.

including those from the Boost Sample, have a young person cross-sectional weight.

The longitudinal weights may be used for analyses of more than one sweep of data. Young people who responded to sweeps 11 and 12, after their main carer responded to every previous sweep of GUS, have a young person longitudinal weight.

The application of longitudinal and cross-sectional weights for analysis of multiple sweeps of data is discussed in the GUS weighting review published alongside the sweep 10 data⁵ (Lubian, Tipping and Bradshaw, 2021).

Table 4.2 Description of weight variables in the data file

Variable name	Label
DkWTchld	Sw12 Young person longitudinal weight
DkWTchd2	Sw12 Young person cross-sectional weight

⁵ Lubian, K., Tipping, S. and Bradshaw, P. (2021) Survey weights and longitudinal analysis – Summary findings from a Growing Up in Scotland Working Paper, Edinburgh: Scottish Centre for Social Research. Access at: [GUS weighting review \(PDF\)](#).

5 Using the data

The GUS sweep 12 data collected from cohort members consists of the following SPSS file:

GUS_SW12_B.sav	2206 cases	Birth Cohort 1
----------------	------------	----------------

5.1 Variables on the data file

The data file contains questionnaire variables (excluding variables used for administrative purposes) and derived variables. The variables included in the file are detailed in the variable list in the data documentation. As far as possible they are grouped in the order they were asked in the interview/questionnaire.

Please note that variable labels in the variable list cannot be relied upon to capture the detail of the question wording, or the answer categories used. For the precise question wording, please refer to the interview documentation.

Data from the CAPI and CAWI modes has been merged into a single dataset. The variable DIMode enables identification of the completion mode. There were no notable differences between questions asked across the different modes.

As noted above, in a few cases, participants provided the same data twice – completing both the CAWI and, later, interviewer-led versions of a questionnaire. Where this was the case, only one set of data was retained for inclusion. In all cases, this was the in-home/CAPI/CASI version because it would be more recent. Furthermore, data provided via the CAPI approach was more likely to be complete (i.e. all questions had been answered) than that provided via CAWI.

For variables with answers following a scale, such as ‘Strongly agree’ to ‘Strongly disagree’ for instance, it must be noted that the order of the answer categories may not follow systematically an ascending or descending scale throughout the list of variables. Also, the answers may equally refer to positive or negative statements. The phrasing of the question and the list of answers provided on the showcards – if any – shape the variables. The user must therefore take these variations into account when creating derived variables.

5.2 Variable naming convention

Variable names are normally made up of 8 characters. The first indicates the source of the variable, the second the year of collection and the rest is an indication of the question topic. Therefore, where the same question was asked in the different sweeps, the names will usually be the same apart from the second character. If a variable name has changed substantially between sweeps this is marked in the variable list.

The naming convention is summarised in Table 5.1.

Table 5.1 GUS variable naming conventions – BC1			
Character no.			
1		2	
Source of data		Sweep	
Non-sequential capitals: A, D, M, P, C		Sequential lower case: a, b, c, etc...	
Source code	Details	Sweep code	Sweep no./Child's age
AL	Area level variable	a	1/10 months
D	Derived variable	b	2/Almost 2 years
DP	Derived variable from partner int	c	3/Almost 3 years
DWP	DWP variable	d	4/Almost 4 years
M	Main carer/adult interview	e	5/Almost 5 years
P	Partner interview	f	6/Almost 6 years
C	Cohort member (young person) interview	g	7/Almost 8 years
W	Young person height/weight	h	8/Around 10-11 years (in Primary 6)
Z	Z-score variable	i	9/Around 12-13 years (in Secondary 1)
		j	10/Around 14-15 years (in Secondary 3)
		k	11/Around 17-18 years (in Secondary 6)
		l	12/Around 19-20 years

5.3 Variable labels

In the sweep 12 dataset the variable labels have been shortened to around 40 characters as far as possible. Although the labels give an indication of the topic of the question **it is essential to refer to the questionnaire documentation to see the full text of the question and the routing applied to that variable.**

For example, for CIPolaca_9 the label in the dataset has been reduced to 'something else'. However, the questionnaire illustrates that the full response available to the participant was 'None of these but have done something else'.

5.4 Derived variables

Derived variables included in the dataset are listed with the questionnaire variables for the same topic. The SPSS syntax used to create them is available on request by contacting the GUS research team on gus@scotcen.org.uk.

5.5 Multicoded questions

Some questions in the survey enabled participants to give more than one answer. In the dataset, each of the answer options for these questions has been converted into a binary variable, with the people who selected that option coded 1 and the rest coded 0.

5.6 Indicators and summary variables

5.6.1 Household details collected at sweep 12

Questions on the cohort member's gender, age, employment, income, and education were included in the questionnaire for all participants. Details for the main carer and (where applicable) the main carer's resident partner were not collected at this sweep beyond any updates to their contact details.

In a small number of cases, the cohort member ended the interview before answering some of these questions and these details were therefore not collected at sweep 12. For proxies for these details, users are referred to the GUS BC1 sweep 11 or prior datasets which are also available through the UK Data Service.

5.6.2 Household data

Contrary to previous sweeps, only the cohort member was interviewed at sweep 12 and so the demographic details of the rest of the household were not collected. Contact details were collected for the cohort member, main carer, and (where applicable) the main carer's resident partner, along with details of the relationship between the cohort member and the main carer and resident partner.

5.6.3 National Statistics Socio-economic Classification (NS-SEC)

The National Statistics Socio-economic Classification (NS-SEC) is a social classification system that attempts to classify groups of people on the basis of employment relations, based on characteristics such as career prospects, autonomy, mode of payment and period of notice. There are fourteen operational categories representing different groups of occupations (for example higher and lower managerial, higher and lower professional) and a further three 'residual' categories for full-time students, occupations that cannot be classified due to a lack of information or other reasons. The operational categories may be collapsed to form a nine, eight, five or three category system.

The sweep 12 dataset includes the five-category system in which cohort members are classified as managerial and professional, intermediate, small employers and own account workers, lower supervisory and technical, and semi-routine and routine occupations. A sixth category 'never worked' is also coded on this variable. The decision on whether or not this category should be included as a separate category, incorporated with category 5 'semi-routine or routine' or set to 'missing' is dependent on the particular analysis to which it is being applied.

Further information on NS-SEC is available from the National Statistics website: [SOC 2020 Volume 3: the National Statistics Socio-economic Classification \(NS-SEC rebased on the SOC 2020\) - Office for National Statistics \(ons.gov.uk\)](#)

Cohort member employment details collected at sweep 12 were used to classify their employment details using the 'self-coded' method of deriving NS-SEC (CINsecoc).

5.6.4 Area-level variables

Scottish Government Urban/Rural Classification

The dataset includes a binary measure of urban/rural location (ALKurin2). This is based on the Scottish Government's two-fold urban rural classification which is itself derived from the more detailed six-fold classification shown in Table 5.1.

The Scottish Government Urban Rural Classification was first released in 2000 and is consistent with the Government's core definition of rurality which defines settlements of 3,000 or less people to be rural. It also classifies areas as remote based on drive times from settlements of 10,000 or more people. The definitions of urban and rural areas underlying the classification are unchanged.

Table 5.1 Scottish Government Six-fold and Two-fold Urban Rural Classifications

Classification	Description – six-fold	Description – two-fold
1. Large Urban Areas	Settlements of over 125,000 people	1. Urban
2. Other Urban Areas	Settlements of 10,000 to 125,000 people	1. Urban
3. Accessible Small Towns	Settlements of between 3,000 and 10,000 people and within 30 minutes' drive of a settlement of 10,000 or more	1. Urban
4. Remote Small Towns	Settlements of between 3,000 and 10,000 people and with a drive time of over 30 minutes to a settlement of 10,000 or more	1. Urban
5. Accessible Rural	Settlements of less than 3,000 people and within 30 minutes' drive of a settlement of 10,000 or more	2. Rural
6. Remote Rural	Settlements of less than 3,000 people and with a drive time of over 30 minutes to a settlement of 10,000 or more	2. Rural

For further details on the classification see the Scottish Government's website: [Scottish Government Urban Rural Classification 2020 - gov.scot \(www.gov.scot\)](https://www.gov.scot/publications/urban-rural-classification-2020/pages/1-6.aspx). A detailed urban/rural variable with all six categories outlined above is available on request under UKDS Secure Licence.

Scottish Index of Multiple Deprivation

The Scottish Index of Multiple Deprivation (SIMD) identifies small area concentrations of multiple deprivation across Scotland. It is based on a number of indicators – 37 indicators in the 2020 version – in the seven individual domains of Current Income, Employment, Health, Education Skills and Training, Geographic Access to Services (including public transport travel times for the first time), Housing and Crime. SIMD is presented at data zone level, enabling small pockets of deprivation to be identified. The data zones – which for the 2020v2 version have a median population size of 755 – are ranked from most

deprived (1) to least deprived (6976) on the overall SIMD and on each of the individual domains. The result is a comprehensive picture of relative area deprivation across Scotland. The GUS sweep 12 dataset contains SIMD 2020v2. It should be noted that analyses in some GUS reports may be based on earlier versions of SIMD⁶.

In the sweep 12 dataset, the data zones are grouped into quintiles in variable ALksimdq2020. Quintiles are percentiles which divide a distribution into fifths, i.e., the 20th, 40th, 60th and 80th percentiles. Those respondents whose postcode falls into the first quintile are said to live in one of the 20% least deprived areas in Scotland. Those whose postcode falls into the fifth quintile are said to live in one of the 20% most deprived areas in Scotland.

Further details on SIMD can be found on the Scottish Government Website: <https://www.gov.scot/collections/scottish-index-of-multiple-deprivation-2020/>.

Further area-level variables (available via UKDS Secure Licence)

Further geographical measures have been derived and are available through UKDS Secure Licence arrangements. These are outlined below.

Data zones

Data zones are the foundational units for small-area statistics in Scotland providing a common, stable and consistent, small-area geography. The data zone geography covers the whole of Scotland and nests within local authority boundaries. Data zones are groups of output areas based on the 2001 census and have populations of between 500 and 1,000 household residents. Where possible, they have been made to respect physical boundaries and natural communities and, as far as possible, contain households with similar social characteristics⁷.

Intermediate geography

Not all statistics are suitable for release at the data-zone level because of the sensitive nature of the statistics, or for reasons of reliability, and it was apparent that a statistical geography between data zone and local authority was required. The intermediate zones are aggregations of data zones within local authorities and contain between 2,500 and 6,000 people⁸.

Local authority

Local government in Scotland comprises 32 unitary local authorities, responsible for the provision of a range of public services. Local authority areas (also known as council areas) reflect the geographical diversity within Scotland with wide variations in size (from 60 square miles in Dundee City council area to 25,656 square miles in Highland council area) and population (from under 20,000 people in Orkney Islands council area to over 600,000 in Glasgow City council area).

⁶ Information on previous versions of SIMD can be obtained from the Scottish Government website via this link: [Scottish Index of Multiple Deprivation: data and resources - gov.scot](https://www.gov.scot/collections/scottish-index-of-multiple-deprivation-2020/)

⁷ Further information on data zones is available from the Scottish Government Small Area Statistics Guide: [Small area statistics: Data Zones and Intermediate Zones - gov.scot](https://www.gov.scot/collections/scottish-index-of-multiple-deprivation-2020/)

⁸ Further information on intermediate geography is available from the Scottish Government Small Area Statistics Guide: [Small area statistics: Data Zones and Intermediate Zones - gov.scot](https://www.gov.scot/collections/scottish-index-of-multiple-deprivation-2020/)

5.6.5 Strengths and Difficulties Questionnaire (SDQ)

The Strengths and Difficulties Questionnaire (SDQ) is a brief behavioural screening questionnaire designed mainly for use with 4-17-year-olds⁹ but with versions available for those aged 18 and over. The scale includes 25 questions which are used to measure five aspects of the child or young person's development – emotional symptoms, conduct problems, hyperactivity/inattention, peer relationship problems and pro-social behaviour. Further details on the SDQ can be found at sdqinfo.org.

At sweep 12, the hyperactivity and inattention subscale of the age 18 and over SDQ (5 items) was asked of the cohort members as part of the self-completion questionnaire (CAWI or CASI). The full list of SDQ items have been asked of the main carer (about the child cohort member) at most sweeps since the cohort member was aged 3 and were asked of the cohort member themselves at sweeps 10 and 11.

A score can be calculated for difficulties within this subscale, with a higher score indicating greater evidence of difficulties. The dataset includes the constituent items from the cohort member. The composite score of this subscale can be derived with reference to the guidance and syntax provided online at sdqinfo.org. A full list of the constituent variables in the cohort member questionnaire is included in Table 5.2.

Table 5.2 Variables associated with the Strengths and Difficulties Questionnaire

Variable name	Description
CICSDQrt	CI: I am restless, I find it hard to sit down for long.
CICSDQfi	CI: I am constantly fidgeting or squirming.
CICSDQdi	CI: I am easily distracted, I find it difficult to concentrate.
CICSDQth	CI: I think before I do things.
CICSDQwk	CI: I finish the work I'm doing. My attention is good.

5.6.6 Mental wellbeing: selected items from the Students' Life Satisfaction Scale

Life satisfaction is measured through the use of selected general items from the Students' Life Satisfaction Scale¹⁰. These items were asked as part of the self-completion questionnaire for the cohort member and were previously asked at sweeps 7, 9, 10 and 11. Relevant variables are listed in Table 5.3. Note that at sweep 12, not all of the cohort members were students. Whilst the full scale was designed for use with students, as the detail in the table shows, the items included at sweep 12 are general in nature and make no reference to being a student and do not require the participant to be a student to answer them.

⁹ Goodman, R. (1997) "The Strengths and Difficulties Questionnaire: a research note", *Journal of Child Psychology and Psychiatry*, 38, pp581-586.

¹⁰ Huebner, E. S. (1991). Initial Development of the Student's Life Satisfaction Scale. *School Psychology International*, 12(3), 231-240. <https://doi.org/10.1177/0143034391123010>.

Table 5.3 Selected items from the Students' Life Satisfaction Scale

Variable name	Description
CIWed	Do you wish your life was different?
CIWer	Do you feel that your life is just right?
CIWea	Do you feel you have what you want in life?
CIWeg	Do you feel you have a good life?

5.6.7 Anxiety: Generalised Anxiety Disorder Assessment (GAD-2)

The GAD-2¹¹ is a very brief (2 item) self-report scale designed as an initial screening tool for core symptoms of Generalised Anxiety Disorder (GAD). The 2 items are scored 0-3 (total score range 0-6), reflecting the frequency of experiencing core symptoms of GAD in the past 2 weeks. A score of 3 is the preferred cut-point for identifying possible cases, where further diagnostic evaluation for GAD would be warranted in clinical settings¹². The GAD-2 has sound psychometric properties, with good sensitivity and specificity for identifying cases of GAD in primary care.

Questions were asked of the cohort member as part of the self-completion questionnaire and were also included at sweeps 10 and 11.

Table 5.4 GAD-2

Variable name	Description
CIgadNer	In last 2 weeks, bothered by: Feeling nervous, anxious or on edge?
CIgadWoS	In last 2 weeks, bothered by: Not being able to stop or control worrying?

5.6.8 Depression: Selected items from the Patient Health Questionnaire (PHQ-9)

The PHQ-9¹³ includes the full 9-item Depression Module of the self-report Patient Health Questionnaire, designed to screen for depression symptoms in the past two weeks. The items reflect the 9 DSM-IV criteria for depressive disorders. Each item is rated on a 4-point scale (score 0-3, total score range 0-27), reflecting the frequency with which symptoms are experienced. The PHQ-9 has been extensively evaluated and found to be valid as both a severity and diagnostic measure in both patient and community samples. On the full scale, a cut-off score of ≥ 10 has been found to maximise combined sensitivity and specificity overall, and for subgroups, by comparison with semi-structured

¹¹ Spitzer R. L., Kroenke K., Williams J. B. W., Löwe B. (2006). A brief measure for assessing generalized anxiety disorder the GAD-7. Arch. Intern. Med. 166 1092–1097. 10.1001/archinte.166.10.1092.

¹² *Catalogue of Mental Health Measures* (2023). Available at www.cataloguementalhealth.ac.uk

¹³ Kroenke K., Spitzer R. L., Williams J. B. (2001). The PHQ-9: validity of a brief depression severity measure. J. Gen. Intern. Med. 16 606–613. 10.1046/j.1525-1497.2001.016009606.

diagnostic interview assessments of depression. Scores of 10 or higher on the full scale are commonly used to identify individuals with depression¹⁴.

A subset of 8 questions was asked of the cohort member. Variables in the dataset are set out in Table 5.5. On the full, formal scale, items CIPHQMo and CIPHQRe are asked as a single question.

Table 5.5 Patient Health Questionnaire (PHQ-9)	
Variable name	Description
	Over the last two weeks, how often have you been bothered by any of the following problems:
CIPHQSI	Trouble falling or staying asleep, or sleeping too much?
CIPHQTi	Feeling tired or having little energy?
CIPHQAp	Poor appetite or overeating?
CIPHQBa	Feeling bad about yourself - or that you are a failure or have let yourself or your family down?
CIPHQCo	Trouble concentrating on things, such as reading the news or watching television?
CIPHQMo	Moving or speaking so slowly that other people could have noticed?
CIPHQRe	Or the opposite - being so fidgety or restless that you have been moving around a lot more than usual?
CIPHQDe	Thoughts that you would be better off dead, or of hurting yourself in some way?

5.6.9 Alcohol and smoking

As part of their self-completion questionnaire, cohort members were asked a number of questions about alcohol, smoking, and use of other tobacco and nicotine products. Other than items on heated tobacco products, the questions included were adapted from the Health Behaviour in School Aged Children Survey (HBSC¹⁵).

Questions about alcohol and smoking were also asked at sweeps 9, 10 and 11. Therefore, not all cohort members answered all questions about alcohol and smoking at sweep 12 (e.g. 'whether ever had alcoholic drink' was not asked of those who had already reported having an alcoholic drink at sweep 11).

Four new items on use of nicotine pouches and heated tobacco products were included at sweep 12.

Table 5.6 Question items on young person's health behaviours - alcohol and smoking	
Variable name	Description
CISmC	CI Whether ever tried a cigarette
CIBSnC	CI How often smokes now
CIBSeC	CI Whether ever tried e-cigarette or vaping device
CIBESnC	CI How often smoke e-cigarettes or vaping devices now
CINicProd	CI Whether ever tried nicotine pouches
CINicProd2	CI How often uses nicotine pouches now

¹⁴ *Catalogue of Mental Health Measures* (2023). Available at www.cataloguementalhealth.ac.uk

¹⁵ <https://hbcs.org/about/>

CITobac	CI Whether ever tried heated tobacco products
CITobac2	CI How often uses heated tobacco products now
CIBSaC	CI Whether ever had alcoholic drink
CIBaIC	CI How often drank alcohol in the last 30 days
CIALFN	CI How many times have had five or more alcoholic drinks at a time in the last 12 months?

5.6.10 Drug use

Alongside questions on alcohol, smoking, and use of other nicotine products, cohort members were also asked about drug use. These items were adapted from the Longitudinal Study of Australian Children (LSAC¹⁶), Growing Up in Ireland (GUI¹⁷) and the Scottish Adolescent Lifestyle and Substance Use Survey (SALSUS¹⁸). See Table 5.7 below.

Table 5.7 Items on drug use	
Variable name	Description
CIDrugMe	CI Have you ever tried cannabis?
CIDrugMo	CI How often have you used cannabis?
CIDrugOe	CI Have you ever tried any drugs other than cannabis?
CIDrugOo	CI How often have you used drugs other than cannabis?
CIDrugOl..	CI Which drugs other than cannabis have you ever tried?
CIDrugOI_1	Prescription only painkillers
CIDrugOI_2	Cocaine
CIDrugOI_3	Ecstasy
CIDrugOI_4	Poppers
CIDrugOI_5	Tranquilisers
CIDrugOI_6	LSD
CIDrugOI_7	Other

5.6.11 Anti-social behaviour and offending

Questions on anti-social behaviour and offending are mostly adaptations of questions previously asked in the Edinburgh Study of Youth Transitions and Crime¹⁹. These questions are listed in Table 5.8 below. These questions have also been asked at previous sweeps.

Table 5.8 Items on anti-social behaviour and offending - adapted from the Edinburgh Study of Youth Transitions	
Relevant variable names	Description
CIASBany	CI: In last year: done anything or behaved in a way that could be described as antisocial or against the law

¹⁶ <https://growingupinaustralia.gov.au/>

¹⁷ <https://www.growingup.gov.ie/>

¹⁸ <https://www.gov.scot/collections/scottish-schools-adolescent-lifestyle-and-substance-use-survey-salsus/>

¹⁹ <https://www.edinstudy.law.ed.ac.uk/>

CIASBsy	Cl: In last year: how many times taken something from a shop or a store
CIASBry	Cl: In last year: how many times been rowdy or rude in public
CIASBmy	Cl: In last year: how many times stolen money or other things
CIASBof	Cl: In last year: how many times online fraud
CIASBky	Cl: In last year: how many times carried a knife or weapon
CIASBby	Cl: In last year: how many times broken into a locked place to steal something
CIASBdg	Cl: In last year: sold illegal drugs to someone

5.6.12 Parent-Child Attachment: selected items from the People In My Life (PIML) scale

The People In My Life measure is a self-report instrument designed to measure attachment to parents and peers in middle childhood, but many items are relevant at other life stages. The sweep 12 cohort member questionnaire included two items from the Parent Attachment scale. Further information about the PIML scale can be found on the Fast Track Project website:

<https://fasttrackproject.org/measure/people-in-my-life/>

At sweep 12, the cohort member was asked these questions about up to two resident parents and up to one parent living elsewhere. Different participants may have answered each set of questions in relation to different parents depending on who the cohort member was living with and which parent(s) they were in contact with. The particular parent the cohort member is responding about in any set of questions is defined in variables CIPa1Rel, CIPa2Rel and CIPe1Rel.

Please also note that the approaches to identifying parents are also different to those applied at some previous sweeps. Users should consult the relevant questionnaire documentation.

Table 5.9 outlines the relevant variables in the dataset.

Table 5.9 Selected items from People In My Life scale

Variable name	Description
CIPar101, CIPar201, CIPew101C	Cl: ... he/she listens to what I have to say.
CIPar102, CIPar202, CIPew102C	Cl: ... I can count on him/her to help me when I have a problem.

5.6.13 Peer Attachment: selected items from the People In My Life (PIML) scale

In addition to the items on relationships with parents outlined above, the sweep 12 self-completion questionnaire also included selected items from the PIML Peer Attachment Scale. Relevant items are outlined in Table 5.10.

Table 5.10 Selected items from People In My Life Peer Attachment scale	
Variable name	Description
CICrFrl	CI: My friends listen to what I have to say
CICrFrt	CI: I talk to my friends when I am having a problem
CICrFrb	CI: If my friends know something is bothering me, they ask me about it

5.6.14 Additional new measures sourced elsewhere

In addition to the items outlined above, Table 5.11 details items new to GUS at sweep 12 which were either directly sourced elsewhere or adapted from existing sources.

Table 5.11 Additional measures sourced elsewhere		
Variable name(s)	Description	Source(s)
CIDechan, CIDechnu	Living arrangements	Adapted from EOPS-B
CIInstR, CIQUCE	Further and higher education	Adapted from the Millenium Cohort Study (Age 23)
CIAprExpect	Further and higher education	Adapted from the COVID Social Mobility and Opportunities study
CIAprEx, CIZAppPay4	Apprenticeships and training	Adapted from the COVID Social Mobility and Opportunities Study
CIApptl, CIAppleave	Apprenticeships and training	Adapted from the Millenium Cohort Study (Age 23)
CINightS	Paid work and employment	Adapted from the Millenium Cohort Study (Age 23)
CICaStep to CICaPay, CICaswrk, CICaswrb	Paid work and employment	Adapted from the Longitudinal Survey of Young People in England: Cohort 2 (Wave 8 and 9)
CIPvachk to CIPvpaaf	Partner violence	Adapted from the Avon Longitudinal Study of Parents and Children
CIResil	Personality and self-assessment of characteristics	Census of Health and Wellbeing
CIUnsafe, CIUnsaaf	Experiences of discrimination and victimisation	Adapted from Understanding Society Wave 1
CISave	Managing financially	Adapted from Growing Up in Ireland (Age 20)

5.7 Dropped variables

All variables in the questionnaire documentation with '[not in dataset]' next to their name have been deleted from the archived dataset (or have been transformed into derived variables instead).

The following types of variables have been deleted or replaced with a derived variable coded into broader categories in order to reduce the potential to identify individuals:

1. Those containing text
2. Those which contained a personal identifier (e.g. name/address)
3. Those considered to be disclosive, such as:
 - Detailed ethnicity
 - Detailed religion
 - Detailed geography variables
 - Language spoken at home
 - Full interview date
 - Full date of birth
 - Timing variables

There are no geographical variables in the archived dataset beyond a binary area urban-rural classification and the Scottish Index of Multiple Deprivation summary variable. As noted in section 5.6.4, access to more detailed geographic variables is possible via the UKDS Secure Licence facility. Access to variables with text/string content is available on request at the discretion and agreement of Scottish Government as the data controller. Further details on accessing additional data can be found on the study website:

<https://growingupinscotland.org.uk/accessing-gus-data>

5.8 Missing values conventions

The following missing values conventions have been observed:

- 1 Not applicable: Used to signify that a particular variable did not apply to a given respondent, usually because of internal routing
- 8 Don't know/Can't say
- 9 No answer/Refused

5.9 Specific issues to note

A small number of corrections were made in instances where an error occurred either in the questionnaire routing or data collection process. These have been highlighted below.

For case IDnumber 4000169, the face-to-face interview with the cohort member was supported by their parent/carer. At question CIYPAGE (cohort member's age at time of interview), the parent's age was erroneously entered instead of the cohort member's age. The cohort member's age has been derived from their date of birth and the interview date and has replaced the parent/carer's age.

Seven young people answered that they were not registered to vote, at variable CregVot, but that they voted in the UK general election in 2024, at variable CIVoteUKP. This should have been prevented by questionnaire routing. As such, this forms a response anomaly. The data for these variables have been left as they are for users to decide how they are handled.

6 Documentation

The documentation for sweep 12 includes the following:

- Questionnaire (with variable names added)
- List of variables in the dataset
- Showcards (used in the CAPI interviews)
- Survey invitation letters and reminder letters and information leaflet
- Interviewer (project) instructions
- Coding specification

7 Contact details

Further details about the study and a list of publications using the data can be found on the study website: www.growingupinscotland.org.uk

Queries should be directed to the GUS team at the Scottish Centre for Social Research:

gus@scotcen.org.uk

Scottish Centre for Social Research

Scotiabank House

6 South Charlotte Street

Edinburgh, EH2 4AW

T: 0131 240 0210