

Technical notes on the data sources

General Household Survey:¹

The GHS is a multi-purpose annual survey conducted by the national statistical agency, Statistics South Africa (Stats SA), to collect information on a range of topics from households in the country's nine provinces. The survey uses a sample of approximately 30,000 households. These are drawn from census enumeration areas using a two-stage stratified design with probability proportional to size sampling of primary sampling units (PSUs) and systematic sampling of dwelling units from the sampled PSUs. The resulting weighted estimates are representative of all households in South Africa.

The GHS sample consists of households and does not cover other collective institutionalised living-quarters such as boarding schools, orphanages, students' hostels, old-age homes, hospitals, prisons, military barracks and workers' hostels. These exclusions should not have a noticeable impact on the findings in respect of children.

Changes in sample frame and stratification

The sample design for the 2012 GHS was based on a master sample that was originally designed for the Quarterly Labour Force Survey (QLFS) and was used for the GHS for the first time in 2008. The same master sample is shared by the GHS, the QLFS, the Living Conditions Survey and the Income and Expenditure Survey. The previous master sample for the GHS was used for the first time in 2004. This again differed from the master sample used in the first two years of the GHS: 2002 and 2003. Thus there have been three different sampling frames during the 11-year history of the annual GHS, with the changes occurring in 2004 and 2008. In addition, there have been changes in the method of stratification over the years. These changes could compromise comparability across iterations of the survey to some extent, although it is common practice to use the GHS for longitudinal monitoring and many of the official trend analyses are drawn from this survey.

Weights

Person and household weights are provided by Stats SA and are applied in Children Count analyses to give estimates at the provincial and national levels.

The GHS weights are derived from Stats SA's mid-year population estimates. The population estimates are revised retrospectively from time to time when it is possible to calibrate the population model to larger population surveys (such as the Community Survey) or to census data. In 2013, Stats SA revised the population model to produce mid-year population estimates in light of the census 2011 results. The new data were used to adjust the benchmarking for all previous GHS data sets, which were re-released with the revised population weights by Stats SA.² All the Children Count indicators have been re-analysed retrospectively, using the revised weights provided by Stats SA. The estimates are therefore comparable over the period 2002 to 2012. The revised weights particularly affected estimates for the years 2002 – 2007. Users may find that the baseline estimates reported here are different from those reported in previous editions of the *South African Child Gauge*. The revised indicators for all the intervening years are available on the website: www.childrencount.ci.org.za.

Survey data are prone to sampling and reporting error. Some of the errors are difficult to estimate, while others can be identified. One way of checking for errors is by comparing the survey results with trusted estimates from elsewhere. Such a comparison can give an estimate of the robustness of the survey estimates. For this project, weighted GHS population numbers were compared with population projections from the Actuarial Society of South Africa's ASSA2008 AIDS and Demographic model (full version).

Analyses of the 11 surveys from 2002 to 2012 suggest that some over- and under-estimation may have occurred in the weighting process, but that much of this has been resolved by the adjusted weights produced by Stats SA:

- When comparing the previously weighted 2002 data with the ASSA2008 AIDS and Demographic model estimates, the number of children appeared to have been under-estimated by 5% overall, with the most severe under-estimation in the youngest age group (0 – 4 years) where the weighted numbers of boys and girls yield under-estimations of 15% and 16% respectively. Running the same comparison against the GHS2002 but applying the revised weights, the difference is substantially reduced. Assuming that ASSA2008 provides the best population estimates, the newly-weighted GHS2002 is much more closely aligned in all age groups, with the 0 – 4-year group slightly under-estimated by 2%. The next age group (5 – 9 years) was previously under-estimated for both boys and girls, at around 7% each. The revised weights produce a slight over-estimate in the region of 4% of boys and 3% for girls in this age group. The 10 – 14-year age group also appears to have been slightly over-estimated when using revised weights (around 3% for both boys and girls), whereas the previous weights had produced a slight under-estimate for this age group. In contrast, the previously weighted data yielded over-estimates of boys and girls in the upper age group (15 – 19 years), with the GHS over-counting these children by about 5%. After re-weighting, the GHS appears to slightly under-count this age group by nearly 6% for boys, and a lower 2.5% for girls. Over the entire 0 – 19-years age group, the population numbers yielded by the revised GHS weights deviate by less than 1% from the ASSA2008 projections. Overall, then, it appears that the re-weighting of the GHS has produced more plausible estimates for 2002, which is the baseline year for most Children Count analyses.
- Population weights derived from the GHS 2012 data are similarly close to the ASSA2008 projects for the 0 – 19-year age group overall, with a difference of less than 1% between the two estimates for both boys and girls. A detailed comparison of individual age groups has not been undertaken as it is acknowledged that the ASSA2008 model was likely to have over-estimated the extent of AIDS mortality and mother-to-child transmission of HIV in South Africa.³ This in turn is likely to affect the child population projections in more recent years. A new model is under development and will address these problems, but was not available at the time of publication.

Apparent discrepancies in male-to-female ratios over the 11 years of data may slightly affect the accuracy of the Children Count estimates. From 2005 to 2008, consistently distorted male-to-female ratios mean that the total estimates for certain characteristics would be somewhat slanted toward the male pattern. This effect is reduced from 2009, where more even ratios are produced, in line with the modelled estimates.

Reporting error

Error may be present due to the methodology used, i.e. the questionnaire is administered to only one respondent in the household who is expected to provide information about all other members of the household. Not all respondents will have accurate information about all children in the household. In instances where the respondent did not or could not provide an answer, this was recorded as "unspecified" (no response) or "don't know" (the respondent stated that they didn't know the answer).

SOCPEN database⁴

Information on social grants is derived from the Social Pensions (SOCPEN) national database maintained by the South African Social Security Agency (SASSA), which was established in 2004 to disburse social grants for the Department of Social Development. Prior to this, the administration of social grants and maintenance of the SOCPEN database was managed directly by the department and its provincial counterparts.

There has never been a published, systematic review of the social grants database, and the limitations in terms of validity or reliability of the data have not been quantified. However, this database is regularly used by the department and other government bodies to monitor grant take-up, and the computerised system, which records every application and grant payment, minimises the possibility of human error. Take-up data and selected reports are available from the department on request throughout the year. Children Count provides grant take-up figures as at the end of March.

References

- 1 Statistics South Africa (2003 – 2013) *General Household Survey Metadata 2002 – 2012*. Pretoria: Stats SA. Available: <http://interactive.statssa.gov.za:8282/webview/>.
- 2 Statistics South Africa (2013) *General Household Survey 2012. Statistical Release P0318*. Pretoria: Stats SA.
- 3 Actuarial Society of South Africa (2012) *Note to users of the ASSA2008 AIDS and Demographic model, 27 August 2012*. <http://aids.actuarialsociety.org.za/>
- 4 South African Social Security Agency (2004 – 2013) SOCPEN social grants data. Pretoria: SASSA.