Sampling-Weighting Design and Wealth index of The Statistical Survey on Functioning and Foundational Learning Skills of Children Living in Georgian Households

Dr. Giorgi Mikeladze

Ivane Javakhishvili Tbilisi State University, Georgia

Abstract

In 2024, the statistical survey was conducted to assess the well-being and foundational learning skills of children living in households in Georgia. The 2014 General Population Census database of Georgia was used for creation sample and statistical weight calculations. The target group of the statistical survey consisted of resident children aged 7–14 in Georgia. To achieve the statistical objectives of the survey, a stratified three-stage cluster random sampling design was implemented. The primary sampling unit (PSU) was the enumeration area of the census; the secondary sampling unit (SSU) was the household address; and the tertiary sampling unit (TSU) was all children aged 7–14 living in the selected household. The stratification variables were defined as the region and type of settlement, totaling 21 strata. To ensure a 2.5% relative margin of error, which was calculated based on a modified version of the standard error formula for complex surveys that accounts for the design effect of cluster sampling and adjustments in statistical weights across strata (Lohr, 2021), total sample size was determined of 1,200 households. Based on the variation by settlement type, it was decided that the number of households to be surveyed per cluster would be 5 in urban areas and 8 in rural areas. At the initial stage, the sample was distributed across regions proportionally to the square root of the number of households in each region. According distribution, the number of clusters was determined to be 193. In the next stage, the clusters were allocated across selected districts by settlement type, and based on this, the final sample size amounted to 1,199 households. To achieve the pre-defined response target of the survey, each fieldwork team was assigned three times more addresses than needed, bringing the final gross sample size to 3,597 households. Data were collected from 2,047 households, yielding a response rate of 56.9%. To reach the required number of respondents (children aged 7–14), additionally, data collected in 567 households using the snowball sampling method. For household weight calculation, probabilities of cluster selection, household selection within the cluster, and response rates were analyzed. Additionally, individual weights were calculated based on clusters. It is

noteworthy that, in line with UNICEF methodology, statistical weights were standardized at the stratum level. UNICEF's Multiple Indicator Cluster Survey (MICS) methodology also includes the calculation of a wealth index. To calculate this index, variables indicating material wealth of the household were selected at: 1) Urban level, 2) Rural level, and 3) National level. Subsequently, factor analysis was conducted, resulting in the identification of factors: one at the urban-rural level, the other at the national level. Regression analysis was carried out using the factors and a binary variable representing settlement type. Based on the predicted values of the dependent variable from this empirical model, quintile groups for the wealth index were calculated. Weighted values were used in both the factor and regression analyses, and the wealth index was calculated at the individual (child) level, not at the household level.

Keywords: sampling design, statistical weights, functioning and foundational learning skills, wealth index, relative marginal error

Acknowledgments: *This work was supported by Shota Rustaveli National Science Foundation of Georgia (SRNSFG) Project №FR-22-25281*