NEPS
National Educational Panel Study

Research Data

Anne Konrad, Ariane Würbach

Samples, Weights and Nonresponse

NEPS Starting Cohort 8 — Grade 5 (2022)

Education for the World of Tomorrow

Wave 2

LIfBi
LEIBNIZ INSTITUTE FOR
EDUCATIONAL TRAJECTORIES

Leibniz Institute for Educational Trajectories (LIfBi)
Wilhelmsplatz 3, 96047 Bamberg
Director: Prof. Dr. Cordula Artelt
Administrative Director: Dr. Stefan Echinger
Bamberg; April 15, 2026

Samples, Weights, and Nonresponse: NEPS Starting Cohort 8 – Grade 6 (2023). Education for the World of Tomorrow. Wave 2

Anne Konrad, Ariane Würbach

Leibniz Institute for Educational Trajectories

NEPS Technical Report for Weighting, referring to DOI:10.5157/NEPS:SC8:2.0.0

E-mail address:

statistik@lifbi.de

Samples, Weights, and Nonresponse: NEPS Starting Cohort 8 – Grade 6 (2023). Education for the World of Tomorrow. Wave 2

Abstract

With the study *Education for the World of Tomorrow*, Starting Cohort 8 (SC8), the National Educational Panel Study (NEPS) has launched a second cohort in lower secondary level from grade 5 onwards. The first cohort in grade 5 was Starting Cohort 3 (SC3), twelve years ago. To ensure comparability with SC3, the sampling design of SC8 was largely conducted analogously that of SC3. Both SC3 and SC8 respect different timings in transitions in lower secondary education, attributable to different educational systems in the Federal States of Germany. Some Federal States in Germany educate students in primary schools from Grade 1 to Grade 6, whereas the majority of primary schools educate students from Grade 1 to Grade 4. The transition to lower secondary education is also decoupled from primary and lower secondary education in some Federal States. These Federal States provide education to students in schools only covering Grade 5 and 6 (so called *Orientierungsstufe*). The corresponding grade 5 students will leave their institutional context in which they were originally sampled. To compensate this early loss of students, a refreshment sample of students in grade 7 will be established in the schools already sampled in grade 5 and by augmenting schools in the respective Federal States, in particular.

This report provides details on the provided weights in wave 2.

Keywords

NEPS SC8, stratified two-stage sampling, unit nonresponse, weighting, calibration

1 Introduction

This weighting report complements the previous weighting report for wave 1 (Konrad et al., 2025) and provides details on wave 2 of Starting Cohort 8 (SC8) of the National Educational Panel Study (NEPS). It refers to the Scientific Use File (SUF; DOI:10.5157/NEPS:SC8:2.0.0). Starting Cohort 8 is the second cohort focusing on students in grade 5 and their pathways through lower secondary education, following SC3, which was initiated twelve years earlier.[1] A sample of students was drawn from both regular schools and special-needs schools. Regular schools comprise all *allgemeinbildende Schulen*, that are, schools of general education according to the definition of the Kultusministerkonferenz (2012). Special-needs schools were restricted to those with a focus on learning.

For the provision of weights, the sampling design as well as the different processes leading to participation decisions at both the institutional and individual levels must be taken into account. Several particularities have to be considered in the weighting process. The aim of this report is to provide detailed information on the computation of the weights for wave 2. Further details on the sampling design, the initial nonresponse adjustment at the school and student levels, and the final calibration can be found in Konrad et al. (2025).

The remainder of this report referring to the second Scientific Use File (SUF) of SC8 Version 2.0.0 (DOI:10.5157/NEPS:SC8:2.0.0) is as follows. Section 2 describes the nonresponse adjustment in wave 2 using response propensity models and response homogeneity groups. Section 3 explains the construction of cross-sectional and longitudinal weights. The computation of user-defined longitudinal weights for SC8 is discussed in Section 4. Finally, Section 5 concludes with a brief discussion of the weights provided.

2 Nonresponse adjustment in wave 2

In general, nonresponse adjustment inflates the initial weights to compensate for the reduction in sample size due to panel attrition and to mitigate potential nonresponse bias. These methods rely on auxiliary information that is available for both respondents and nonrespondents. In wave 2, auxiliary information on nonrespondents is available from wave 1, which enables the use of response propensity score weighting.

2.1 Propensity score weighting

The response propensity, $\hat{p}_i = P(r_i = 1 \mid x_i)$, describes the probability that a sampled unit i participates in the survey, conditional on a set of auxiliary variables x_i observed for both respondents and nonrespondents. The variable r_i is an indicator function with $r_i = 1$ if unit i responses, and zero otherwise.

A common parametric approach modeling response propensities is the probit model,

$$\hat{p}_i = \Phi(x_i^\top \beta), \tag{1}$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution.

[1] For more specific information on research topics in the NEPS, see Blossfeld and Roßbach (2019).

For the nonresponse adjustment, a mixed probit model with institution as a random effect is applied. Table 3 reports the estimation results of the model. Individual tracking and native language significantly decrease the probability of participation. In contrast, school types such as Gymnasium, integrated comprehensive schools, and Orientierungsstufe, as well as participation in the previous wave, significantly increase the response probabilities.

A critical issue with parametric response propensity models is that they may yield very small estimated response probabilities for some units. This can result in extremely large weights and, consequently, unstable survey estimates.

A potential remedy is to group units into response homogeneity groups (RHGs) and assign a common adjustment factor within each group, rather than relying on explicit individual response probabilities.

2.2 Response homogeneity groups

Instead of using individual estimated response probabilities, sampled units with similar response behavior can be grouped, and nonresponse adjustment can be performed at the group level (Eltinge & Yansaneh, 1997). These groups are commonly referred to as response homogeneity groups (RHGs).

A widely used approach to constructing RHGs is the score method (Eltinge & Yansaneh, 1997; Little, 1986), which consists of three main steps. First, response probabilities $\hat{p}_i$ are estimated for all sampled units $i \in s$ using the parametric probit model in (1). Second, sampled units are ordered according to their estimated response probabilities. Third, the ordered sample is partitioned into G groups of approximately equal size, resulting in groups that are homogeneous with respect to response propensities.

Within each response homogeneity group $g = 1, \ldots, G$, a common nonresponse adjustment factor is applied to all units. The adjustment factor for group g is defined as

$$\hat{p}_g = \frac{\sum_{i \in r_g} \hat{p}_i}{n_g}, \tag{2}$$

where n_g denotes the number of sampled units in group g and r_g the number of respondents in that group. $G = 10$ response homogeneity groups are used.

3 Computation of cross-sectional and longitudinal weights in later waves

In longitudinal panel studies, the weighting process requires additional weighting steps to account for participation in later waves, beyond the three basic steps of computing design weights, nonresponse adjustment, and calibration (see Konrad et al. (2025) for details). The data have both a cross-sectional and a longitudinal dimension, and therefore two sets of weights are provided. Cross-sectional weights refer to the target population of a specific wave and should be used when the research interest focuses on the analysis of a single wave. Longitudinal weights, in contrast, are related to units that participate in multiple consecutive waves and are appropriate when the interest lies in analyzing changes over time for the same units.

In general, cross-sectional weights are computed recursively as

$$\begin{aligned} w_1 &= d_1 \cdot p_1^{-1} \cdot g_1, \\ w_2 &= d_1 \cdot p_1^{-1} \cdot p_2^{-1} \cdot g_2, \\ &\vdots \\ w_T &= d_1 \cdot \prod_{t=1}^{T} p_t^{-1} \cdot g_T, \end{aligned} \tag{3}$$

where d_1 denotes the design weight from wave 1, p_t the response propensity in wave t, and g_t the calibration factor in wave t. As in wave 1, a linear distance function is used for calibration, resulting in the GREG estimator, with school branches and gender included as auxiliary variable.Table 4 shows the distributions of school branches, gender and federal states in the panel and among participants in wave 2 for Grade 2, compared to the corresponding distributions in the 2023/2024 school frame. The weighted sample distributions of gender and school branches match those in the frame as as result of the calibration.

Due to the lack of information on nonrespodents in some federal states in wave 1, it was not possible to model response probabilities p_1 as described in Section 2.1. Instead a two-step calibration approach was applied (see Konrad et al. (2025) for details). Consequently, no individual response probabilities are available that could be used to compute the wave 2 weights according to (3). Instead, we approximate the response probability by a constant $\alpha = n_1/n_2$, which corresponds to the intercept from a model with only an intercept estimated in wave 1. The weights in wave 2 are therefore computed as

$$w_{i2} = \begin{cases} d_i \cdot \alpha^{-1} \cdot p_2^{-1} & \text{for } i \in r_2 \cap r_1, \\ d_i \cdot (1-\alpha)^{-1} \cdot p_2^{-1} & \text{for } i \in r_2 \setminus r_1, \end{cases} \tag{4}$$

$i \in r_2 \setminus r_1$ are temporary dropouts, i.e., individuals who were nonrespondents in wave 1 but participated in wave 2. In wave 2, there are 262 such temporary dropouts.

Longitudinal weights are also computed recursively

$$\begin{aligned} w_{12} &= d_1 \cdot p_1^{-1} \cdot g_1 \cdot p_2^{-1} = w_1 \cdot p_2^{-1}, \\ &\vdots \\ w_{1T} &= d_1 \cdot p_1^{-1} \cdot g_1 \cdot \prod_{t=2}^{T} p_t^{-1} = w_1 \cdot \prod_{t=2}^{T} p_t^{-1}. \end{aligned} \tag{5}$$

In the corresponding Scientific Use File, longitudinal weights are only provided for units that participate continuously from the first wave onwards. For longitudinal analyses starting at later waves, users need to construct appropriate longitudinal weights themselves (see Section 4).

4 Computation of own longitudinal weights

Starting with this wave, the Scientific Use File (SUF) provides, in addition to cross-sectional and longitudinal weights, the estimated response propensities for each wave. This enables users

to construct longitudinal weights for any pair of waves by themselves. Basis for the computation for this are the formulas given in (5). For example, if the analysis focuses on units that participate continuously from wave 2 to 4, the corresponding longitudinal weight is given by

$$w_{24} = w_2 \cdot p_3^{-1} \cdot p_4^{-1}. \tag{6}$$

That means the starting point is always the cross-sectional weight from the respective start wave which is multiplied by the response probability for each following wave taken into consideration.

In general, longitudinal weights for a sequence of waves are obtained by starting from the cross-sectional weight of the initial wave and successively multiplying the inverse of the response propensities for all each subsequent waves.

5 Conclusion

This paper provides an overview on the derivation of weights for wave 2. Sampling design and nonresponse were adjusted by calibration to the school frame 2023/2024.

Table 1 summarizes the design information and the different weights in the SUF release version DOI:10.5157/NEPS:SC8:2.0.0. Besides target (student) (`ID_t`) and institutional (school) (`ID_i`) identifiers, design information for the entire cohort is available.[2] This information covers the sample at start[3], the explicit sampling stratum (`stratum_exp`) as well as the implicit sampling strata (`stratum_imp`).

Weights at the institutional (`w_i`) and the target (`w_t_std`) level are given for the entire cohort (initial participants).[4] For all participants in a particular wave, cross-sectional weights (e.g., `w_t1`) are published. With respect to panel progress, longitudinal weights are also available (e.g., `w_t1to2`). Starting in wave 2, we also provide wave-specific response probabilities for own calculations (e.g., `p_t2`).

In SC8, no trimming was performed because it has the potential to introduce bias, undo the effects of the previous weighting steps and may be ineffective (Chen et al., 2017; Kimberly & Valliant, 2012). To prevent extreme values in the weights, we reviewed the weights at each step and, if necessary, combined groups.

All weights apart from the institutional weight are provided in a standardized form. Standardized weights have mean one and sum up to the number of participants. Summary statistics for all kind of weights provided are given in Table 5.

For further information on weighting please contact `statistik@lifbi.de`.

[2] Due to data protection, the design information is not available in the download version of the SUF.

[3] Currently, all students were sampled in grade 5. However, in wave 3 the sample is going to be augmented with new schools and refreshed by new students from the schools originally sampled in grade 5.

[4] The institutional weight as well as the explicit and implicit stratification variables belong to the institution and thus are equal for all cases within the institution.

Acknowledgements

This paper uses data from the National Educational Panel Study (NEPS; see Blossfeld and Roßbach, 2019). The NEPS is carried out by the Leibniz Institute for Educational Trajectories (LIfBi, Germany) in cooperation with a nationwide network.

6 References

Blossfeld, H.-P., & Roßbach, H. G. (Eds.). (2019). *Education as a lifelong process: The German National Educational Panel Study (NEPS). Edition ZfE* (2nd). Springer VS.

Chen, Q., Elliot, M. R., Haziza, D., Yang, Y., Gosh, M., Little, R. J., Sedransk, J., & Thompson, M. (2017). Approaches to improving survey-weighted estimates. *Statistical Science*, *32*(2), 227–248.

Eltinge, J. L., & Yansaneh, I. S. (1997). Diagnostics for formation of nonresponse adjustment cells, with an application to income nonresponse in the us consumer expenditure survey. *Survey methodology*, *23*, 33–40.

Kimberly, H., & Valliant, R. (2012). Comparing alternative weight adjustment methods. *Proceedings of the Section on Survey Research Methods*, 4696–4710.

Konrad, A., Würbach, A., & Aßmann, C. (2025). *Samples, Weights, and Nonresponse: NEPS Starting Cohort 8 - Grade 5 (2022). Education for the World of Tomorrow. Wave 1*. Leibniz Institute for Educational Trajectories. Bamberg. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC8/1-0-0/NEPS_SC8_Weights_1-0-0_en.pdf

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

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/

Appendix

6.1 Tables

Table 1: Variables included in the weighting data set for SC8 SUF version 2.0.0

Variable	Applies to	Content
ID_i	270	Identifier for the school the target person was initially sampled in
ID_t	6141	Identifier for target person
Design information		
sample	6141	Part of the sample the target person belongs to
stratum_exp	6141	Explicit stratum referring to the school
stratum_imp1_R	6141	Implicit stratum (Federal State the school is located in according to sampling frame)
stratum_imp2_R	6141	Implicit stratum (regional classification according to sampling frame)
stratum_imp3_R	6141	Implicit stratum (funding according to sampling frame)
Final weights for initial participants		
w_i	6141	Final weight for institution, calibrated
w_t	6141	Final weight for target, calibrated
w_t_std	6141	Final weight for target, calibrated and standardized
Final weights for wave-specific participants		
w_t1	5763	Cross-sectional weight for targets participating in Wave 1, calibrated
w_t1_std	5763	Cross-sectional weight for targets participating in Wave 1, calibrated and standardized
w_t2	5212	Cross-sectional weight for targets participating in Wave 2, calibrated
w_t2_std	5212	Cross-sectional weight for targets participating in Wave 2, calibrated and standardized
w_t1to2_std	4950	Longitudinal weight for targets participating in Wave 1 to 2, standardized
Wave-specific response probabilities		
p_2	5212	Response probability in Wave 2

Table 2: Panel progress of SC8 by wave.

			Panel Cohort		Status at the end of the wave			
Wave (Time)	Study number	Total size	Not used	Used sample	Participants	Temporary dropout	Final dropout (in wave)	Final dropout (after wave)
1 2022/2023	Grade 5	6141	0	6141	5763	378	—	21
2 2023/2024	Grade 6	6120	0	6120	5218	793	109	157

n.a. - “not (yet) available”.

Table 3: Models estimating the individual participation propensities for students in Wave 2 of SC8 used to derive adjustment factors for adjusted wave-specific cross-sectional and longitudinal weights.

	Wave 2
Constant	0.504^{***}
	(0.189)
School branch: GY	0.524^{***}
	(0.191)
School branch: HS	0.144
	(0.230)
School branch: IG	0.448^{**}
	(0.208)
School branch: MB	0.138
	(0.215)
School branch: OS	0.539^{*}
	(0.299)
School branch: RS	0.413
	(0.254)
Age group: younger half	0.017
	(0.049)
Native language: missing	0.009
	(0.127)
Native language: other	-0.257^{***}
	(0.081)
Gender: female	-0.045
	(0.048)
Student participated in Wave 1	0.548^{***}
	(0.087)
Individual tracking: yes, school refusal	-1.185^{***}
	(0.209)
Individual tracking: yes	-1.878^{***}
	(0.116)
Observations	6,120

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$; standard errors are given in parentheses. To model individual participation with random effect model the `glmer` function with a probit link provided in R (R Core Team, 2020) was used. Reference categories are: School branch (FS), Age group (older half), Native language (German), Gender (male), Student participated in Wave t (no), Individual tracking (no, i.e., still in school context).

Table 4: Distributions in Grade 6 at student level.

	Frame		Panel		Part. wave 2		
	abs.	**%**	**abs.**	**%**	**abs.**	**unw. %**	**wt. %**
School branch							
FS	17432	2.3	241	3.9	176	3.4	2.3
GY	288878	37.4	3072	50.2	2728	52.3	37.4
HS	52198	6.8	376	6.1	270	5.2	6.8
IG	146817	19	1087	17.8	912	17.5	19
MB	88068	11.4	658	10.8	530	10.2	11.4
OS	59075	7.7	224	3.7	192	3.7	7.7
RS	119574	15.5	462	7.5	404	7.8	15.5
Gender							
male	362172	46.9	3080	50.3	2622	50.3	47
female	409874	53.1	3040	49.7	2590	49.7	53
Federal State							
BB	23786	3.1	149	2.4	122	2.3	5.5
BE	34816	4.5	220	3.6	199	3.8	5.1
BW	104230	13.5	462	7.5	360	6.9	7.7
BY	122626	15.9	745	12.2	597	11.5	13.2
HB	6068	0.8	15	0.2	10	0.2	0.2
HE	56361	7.3	304	5	244	4.7	4.7
HH	16996	2.2	224	3.7	196	3.8	2.4
MV	14949	1.9	50	0.8	43	0.8	0.6
NI	73707	9.5	702	11.5	566	10.9	12.6
NW	168224	21.8	1342	21.9	1207	23.2	18.8
RP	37256	4.8	296	4.8	242	4.6	3.8
SH	26497	3.4	470	7.7	430	8.3	6.7
SL	8493	1.1	5	0.1	5	0.1	0.1
SN	38871	5	252	4.1	216	4.1	4.1
ST	19278	2.5	499	8.2	440	8.4	7.1
TH	19884	2.6	385	6.3	335	6.4	7.2
Sum	**772042**	**100**	**6120**	**100**	**5212**	**100**	**100**

Sources: School frame 2023/2024, A105_T_TR, own data.

Table 5: Summary statistics for all weights provided.

Label of weight	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
w_i	5.1448	21.5891	35.2388	38.6911	50.4531	150.0000	
w_t	24.5004	76.9647	110.1265	123.3254	145.5434	416.3668	
w_t_std	0.1987	0.6241	0.8930	1.0000	1.1802	3.3762	
w_t1	27.2849	80.8199	116.3852	131.4144	152.9027	420.3344	378
w_t1_std	0.2076	0.6150	0.8856	1.0000	1.1635	3.1985	378
w_t2	2.5881	99.5856	127.9059	148.1278	177.6698	1206.3556	929
w_t2_std	0.0175	0.6723	0.8635	1.0000	1.1994	8.1440	929
w_t1to2_std	0.1974	0.5747	0.8850	1.0000	1.1683	7.5898	1191