NEPS
National Educational Panel Study

Research Data

Anne Konrad, Ariane Würbach, Christian Aßmann

Samples, Weights and Nonresponse

NEPS Starting Cohort 8 — Grade 5 (2022)

Education for the World of Tomorrow

Wave 1

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; July 28, 2025

Samples, Weights, and Nonresponse: NEPS Starting Cohort 8 – Grade 5 (2022). Education for the World of Tomorrow. Wave 1

Anne Konrad, Ariane Würbach, Christian Aßmann

Leibniz Institute for Educational Trajectories

NEPS Technical Report for Weighting, referring to DOI:10.5157/NEPS:SC8:1.0.0

E-mail address:

statistik@lifbi.de

Samples, Weights, and Nonresponse: NEPS Starting Cohort 8 – Grade 5 (2022). Education for the World of Tomorrow. Wave 1

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 by augmenting schools in the respective Federal States, in particular.

This report provides details on the sampling design and the provided weights. Based on new insights from the literature, we revised the weighting approach in comparison to that applied in SC3. In the following sections, we provide the rational and description of these changes.

Keywords

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

1 Introduction

Starting Cohort 8 of the National Educational Panel Study (NEPS) is the second cohort that focuses on students in grade 5 and their pathway through lower secondary education, besides the SC3 twelve years ago.[1] A sample of students in regular schools and special-need schools was selected. Regular schools are all “allgemeinbildende Schulen´´, that are, schools of general education according to the definition of the Kultusministerkonferenz (2012). Special-need schools were restricted to those with focus on learning.

For weights provision, the sampling design and the different processes leading to the participation decision at the institutional and the individual level have to be considered. Several particularities have to be considered in the weighting process. The aim of this paper is to provide details on the sampling design, the initial nonresponse adjustments at school and student level, final calibration as well as wave-specific nonresponse adjustments.

The remainder of this report referring to the first Scientific Use File (SUF) of SC8 Version 1.0.0 (DOI:10.5157/NEPS:SC8:1.0.0) is structured as follows: Section 2 gives information on the population and the sampling design to realize SC8. The general weighting process is explained in Section 3. The derivation of the school and student weights is presented in Sections 4 and 5, respectively. Section 6 documents the wave-specific nonresponse adjustments to provide cross-sectional weights. Finally, Section 7 concludes with a brief discussion of the weights provided.

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

2 Population and Sampling

The target population of SC8 includes students in grade 5 in regular and special-need schools offering lower secondary education within the Federal Republic of Germany in the school year 2022/2023. Excluded are students attending schools with a predominant foreign teaching language and students attending special-need schools with another focus than learning. The sampling design of SC8 is very similar to that of the SC3 in 2010, cp. Aßmann et al. (2011), Steinhauer and Zinn (2016), and Steinhauer et al. (2015). Access to the student population was obtained through the schools in which they were enrolled. The initial grade 5 sample was drawn using a two-stage stratified sampling design, with schools selected as primary sampling units (PSU) in the first stage and classes within the selected schools in the second stage. Within the classes, all students were invited for participation in the NEPS, thus constitute the secondary sampling units (SSU). The selection at the first stage was based on a complete list of all regular and special-needs schools in Germany. This school frame contain schools of the general educational system only, e.g. vocational schools are excluded. It was compiled using up-to-date school registers for the school year 2020/2021 that were available from the *Statistical Offices* of all 16 Federal States.

As in SC3, stratified two-stage sampling design was applied with explicit and implicit stratification. For explicit stratification, schools were classified into 7 branches (*Schulzweig*)[2]:

- *Schools leading to upper secondary education and university entrance qualification (Gymnasien, GY)*,
- *Intermediate secondary schools (Realschulen, RS)*,
- *Schools for basic secondary education (Hauptschulen, HS)*,
- *Schools with several courses of education (Schulen mit mehreren Bildungsgängen, MB)*,
- *Comprehensive schools (Integrierte Gesamtschulen, IG) and Rudolf Steiner schools (Freie Waldorfschulen, FW)*,
- *Schools only covering the orientation stage (Schulartunabhängige Orientierungsstufe, OS)*, and
- *Schools offering schooling to students with special educational needs, with focus on learning (Förderschulen, FS)*

(compare variable `stratum_exp` in Table 1). For implicit stratification[3] the following characteristics are used:

- *Federal State* (`stratum_imp1`),
- *Degree of urbanization or regional classification* (`stratum_imp2`), and
- *Sponsorship (public vs. private)* (`stratum_imp3`).

[2] An individual school can offer different school branches. Sampling was based on school branches instead of institutions. Accordingly, schools with several branches were represented several times in the sampling frame.

[3] Sorting the sampling frame by certain characteristics together with a systematic selection is referred to as implicit stratification.

At the first stage, schools were selected by using a systematic probability proportional to size (*pps*) sampling design within each explicit stratum $h = 1, \ldots, 7$. The total number of schools in the population is $M = \sum_{h=1}^{7} M_h$, with M_h as the number of schools in stratum h. From the total list of M =11,402 schools (excluding special-needs schools with a focus other than learning and schools with a non-German language of origin in Hesse), 450 regular schools and 125 special-needs schools were targeted for sampling. The number of sampled schools $m = 575$ were allocated as follows:

- $m_{1=GY} = 101$,
- $m_{2=RS} = 66$,
- $m_{3=HS} = 91$,
- $m_{4=MB} = 71$,
- $m_{5=IG} = 89$ (including *FW*),
- $m_{6=OS} = 32$, and
- $m_{7=FS} = 125$

For the systematic *pps* sampling, we define the measure of size (mos) for school k in stratum h as

$$mos_{kh} = \frac{C^5_{kh}}{\min\{C^5_{kh}; 4\}}, \tag{1}$$

where C^5_{kh} denotes the number of classes in grade 5 that school k in stratum h hosts according to the frame referring to school year 2020/2021. The inclusion probability π_{kh} for school k in stratum h under systematic sampling without replacement is given by

$$\pi_{kh} = m_h \cdot \frac{\frac{C^5_{kh}}{\min\{C^5_{kh};4\}}}{\sum_{k=1}^{M_h} \frac{C^5_{kh}}{\min\{C^5_{kh};4\}}} \tag{2}$$

(Särndal et al., 1992; Valliant et al., 2013).

At the second stage, four classes were randomly selected within each sampled school, provided that the school had at least five classes.[4] Otherwise all available classes were selected. All students of the selected classes are asked to participate. The inclusion probability π_{jkh} for student j in school k in stratum h is calculated from

$$\pi_{jkh} = m_h \cdot \frac{\frac{C^5_{kh}}{\min\{C^5_{kh};4\}}}{\sum_{k=1}^{M_h} \frac{C^5_{kh}}{\min\{C^5_{kh};4\}}} \cdot \frac{\min\{\widetilde{C}^5_{kh}; 4\}}{\widetilde{C}^5_{kh}}, \tag{3}$$

where $\widetilde{C}^5_{jh}$ denotes the number of classes school j in stratum h hosts in school year 2022/2023.

[4] 54 schools decided to participate with all classes.

The design weights for school k and student j are calculated as follows:

$$d_{kh} = \pi_{kh}^{-1} \quad \text{and} \tag{4}$$

$$d_{jkh} = \pi_{jkh}^{-1} \tag{5}$$

with π_{kh} and π_{jkh} defined in (2) and (3), respectively.

3 Weighting

In general, the weighting process involves the following three steps, see Haziza and Beaumont (2017):

Step 1 **Obtain design weights to account for the random sample selection:** The design weights are given by the inverse of the inclusion probability. The inclusion probability describes the probability of each unit in the population to be selected into the sample. The selection process and the derived design weights are described in Section 2.

Step 2 **Nonresponse Adjustment:** The consequences of nonresponse are twofold: First, nonresponse decreases the effective sample size. Second, nonresponse can result in biased estimates if the respondents systematically differ from the nonrespondents. As a remedy, the design weights from step 1 have to be adjusted to compensate for the consequences of nonresponse. Nonresponse adjustment methods include binary regression modelling, response homogeneity groups or calibration. Each of these methods relies on auxiliary information that is available for both respondents and nonrespondents.

Step 3 **Final Calibration:** Often auxiliary information on population totals are available from other reliable sources. For example, the population number of female and male students are known from a school frame. Then, in the calibration step, the weights derived in step 2 are further adjusted such that the survey estimates of female and male students agree with the known number from the school frame.

The weighting process at the school and student levels differ due to variations in accessible information. At school level, we have to omit weighting step 2, because no additional auxiliary information on the nonrespondents is available at the gross level. Nevertheless, the final calibration in step 3 may reduce the potential bias due to nonresponse and undercoverage. The final weights at school level (with i for institutional) are defined by

$$w_{i_k} = d_{kh} g_{kh}^{\text{cal}} , \tag{6}$$

with design weight d_{kh} defined in (4) and g_{kh}^{cal} as final calibration weight (see Section 4).

At student level, in turn, more information on the nonrespondents at gross level is accessible. Therefore, the final weights at student level (with t for target) are given by

$$w_{t_j} = d_{jkh} g_{jkh}^{\text{non}} g_{jkh}^{\text{cal}} , \tag{7}$$

with design weight d_{jkh} defined (5), g^{non} as nonresponse adjustment computed in weighting step 2 (see Section 5.1) and g^{cal} as final calibration weight computed in weighting step 3 (see Section 5.2).

4 Derivation of school weights

Since participation in the NEPS panel study was voluntary for both, the sampled schools and the students, nonresponse can arise at both levels. Two strategies were applied to mitigate nonresponse. First, the NEPS SC8 was supported by the *Ministries of Education* in several Federal States. However, school recruitment still proved to be a particular challenge. Second, if a school of the original sample refused to participate in the NEPS, the loss was compensated for by one of the structurally equivalent replacement institutions drawn in addition to that school. Those replacement schools were schools with identical characteristics in terms of federal state, sponsorship and degree of urbanization as well as similar class sizes. Up to nine replacement schools were drawn for each selected school branch. Altogether, the gross sample included 5,068 school branches (4,478 regular schools and 590 special-needs schools). Steinhauer et al. (2015) give more details on the replacement strategy to prevent bias caused by schools refusal.

Despite the high effort in school recruitment, only 270 schools (thereof 29 special-needs schools) actually agreed to participate in the NEPS, which corresponds to a response rate of 5% at the institutional level, and 47% of the originally planned school sample. Of the 4,798 nonparticipating schools were 12 schools not yet contacted, 3,674 explicitly refused, 1,005 schools just did not respond. The remaining 107 schools that agreed to participate at an earlier stage of recruitment, however, could not afford to participate, mostly due to time or personnel constraints.

We use a calibration approach in order to reduce the potential bias due to nonresponse and undercoverage. The intention of calibration introduced by Deville and Särndal (1992) is to construct a new set of weights by using auxiliary information in order to improve the efficiency of the survey estimates. The design weights are adjusted such that the weighted sample-based estimates for the auxiliary variables coincide with their known population totals available from other reliable sources. For the computation of g^{non}, we apply the minimum distance approach, which minimizes the distance between the initial design and the calibration weight, while still satisfying the calibration constraints. The calibration constraints are defined such that the weighted sampled estimates reproduce the known population totals of the auxiliary variables. Closeness between design and calibration weight is measured by a pre-specified distance function $G(\cdot)$. Let x_k be the vector of auxiliary variables of school k with τ_x as corresponding vector of population totals available from the school frame. Then, the minimization problem at school-level is formalized as

$$\min_{w_{i_k}} \sum_{k=1}^{m} G(w_{i_k}, d_k)$$

subject to the calibration constraints

$$\sum_{k=1}^{m} w_{i_k} x_k = \tau_x \ , \tag{8}$$

with $w_{i_k} = d_k g_k^{\text{cal}}$ as final calibrated weight.[5] We chose a linear distance function which results in the generalized regression (GREG) estimator (Deville & Särndal, 1992; Särndal et al., 1992).

[5] Calibration was performed using the function `calib()` from the R package sampling (version 2.10).

Auxiliary variables at the school-level are school branch and federal state, because these characteristics influence the participation propensities at institutional level, see Table 8.1.[6]

5 Derivation of student weights

In order to be part of the study, one legal guardian of the student had to give written consent. Only students for whom a fully completed consent form was available on the day of the survey were allowed to participate. Out of a total of 20,536 registered students in the selected 5th grade classes (gross sample), 6,141 students (thereof 241 from special-needs schools) were willing to participate and in possession of the required parental declaration of consent (panel sample). See Table 2 for an overview of the case numbers at student level.

At student-level, we differentiate between initial panel participants, i.e. participants with the general willingness to participate in the NEPS panel study, and wave-specific participation, i.e. students that really participate in the respective wave. The weighting processes for initial panel participation is explained in Section 5 and that of wave-specific participation in Section 6.

5.1 Initial Nonresponse Adjustments

Nonresponse adjustment methods rely on auxiliary information available for both respondents and nonrespondents. However, in the first wave less information for the nonrespondents is provided. Moreover, due to data limitations, student-level information is missing for Hesse, Bavaria and Saxony. Because of that limited information, we choose a calibration approach to compute the nonresponse adjustment factor g^{non} (weighting step 2 from above), instead of a logit modelling approach which estimates explicit response probabilities.[7] This proceeding is in line with Kalton and Flores-Cervantes (2003).

At the student-level, the minimization problem is formalized as

$$\min_{\tilde{w}_j^{\text{non}}} \sum_{j=1}^{n} G(\tilde{w}_j^{\text{non}}, d_j)$$

subject to the calibration constraints

$$\sum_{j=1}^{n} \tilde{w}_j^{\text{non}} x_j = \tau_{x^{\text{gross}}} \tag{9}$$

with $\tilde{w}_j^{\text{non}} = d_j g_j^{\text{non}}$ as intermediate calibrated weight and n as the number of students in the sample. In the nonresponse adjustment step, the known total vector $\tau_{x^{\text{gross}}}$ is given at gross level.

Because less auxiliary information on the nonrespondents is available for Bavaria and Saxony, two separate calibrations are conducted for Bavaria and Saxony and the remaining federal

[6] Due to small sample sizes, Saarland was summarized with Rhineland-Palatinate into one category and Bremen was summarized with Lower Saxony into one category for calibration.

[7] The results of the corresponding logit models for panel participation and participation in wave 1 for all federal states, except Bavaria, Saxony and Hesse, can be found in the appendix, see Table 6.

states. In wave 1, the set of auxiliary variables is restricted to the sampling information and delivered information at the gross sample. For Bavaria and Saxony, auxiliary variables are gender and school branch, while for the other federal states, gender, school branch and age group (younger vs. older half)[8] are used.

5.2 Final Calibration

In the final calibration step (weighting step 3 from above), we compute weighting factor g^{cal}. The minimization problem in the final calibration step at school-level is formalized as

$$\min_{w_{t_j}} \sum_{j=1}^{n} G(w_{t_j}, \tilde{w}_j^{\text{non}})$$

subject to the calibration constraints

$$\sum_{j=1}^{n} w_{t_j} x_j = \tau_x \tag{10}$$

with $w_{t_k} = \tilde{w}_j^{\text{non}} g_j^{\text{cal}} = d_j g_j^{\text{non}} g_j^{\text{cal}}$ as final calibrated weight as defined in (7). The calibration is run based on all federal states in common. As auxiliary variable school branch was utilized.

6 Wave-specific participants

Students being part of the SC8 panel can decide in each wave whether they want to participate again or not. We distinguish three different participation statuses in each wave: participant, temporary drop out, and final drop out. A student is considered as final drop out if the panel consent is withdrawn and further participation in the panel is refused. In contrast, a temporary drop out is defined as student who does not participate in the current wave but remains generally willing to continue participation in the panel and has not withdrawn their panel consent. Participants are all students that provide any information.

6.1 Participation in Wave 1

Information for the first survey wave is available for 5,763 students, i.e. they took the tests and/or completed the questionnaire. This represents a participation rate of 93.8%. Students who could not be reached at school on the day of surveying, received a separate self-administered questionnaire via computer-assisted web interview (CAWI). Table 2 gives the number of students and their participation status by wave.

The weighting process to provide cross-sectional weights in wave 1 is exactly the same as that explained in Section 5. The only difference is given by the different number of students. Here we use the 5,763 students that actually attend the survey. Above the basis was the 6,141 students generally willing to participate in the SC8.

[8] Students are categorized into an younger and an older half according to the median age of the entire cohort.

The provided cross-sectional weight for students participating in wave 1 is included in the weighting data as `w_t1`, see Table 1. Longitudinal weights for students participating in all successive waves are published in later waves.

Table 4 compares case numbers from the school frame 2022/2023, the gross sample, the panel sample, and those participating in wave 1.

7 Conclusion

This paper provides an overview on the sampling design for selecting the sample of students in grade 5 in the SC8, and the corresponding derivation of weights. Sampling design and nonresponse were adjusted by calibration to the school frame 2022/2023.

Various kinds of target-specific weights as well as design information are provided. Table 1 summarizes the design information and the different weights in the SUF release version DOI:10.5157/NEPS:SC8:1.0.0. Besides target (student) (`ID_t`) and institutional (school) (`ID_i`) identifiers, design information for the entire cohort is available.[9] This information covers the sample at start[10], 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`) level are given for the entire cohort (initial participants).[11] For all participants in a particular wave, cross-sectional weights (`w_t1`) are published.

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`.

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.

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

[10] Currently, all students were sampled in grade 5. However, in wave 3 the sample is going to be augmented (new students from the schools originally sampled in grade 5) and refreshed (new schools) by grade 7 students.

[11] 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.

8 References

Aßmann, C., Steinhauer, H. W., Kiesl, H., Koch, S., Schönberger, B., Müller-Kuller, A., Rohwer, G., Rässler, S., & Blossfeld, H.-P. (2011). Sampling designs of the National Educational Panel Study: Challenges and solutions. In H.-P. Blossfeld, H. G. Roßbach, & J. von Maurice (Eds.), *Education as a lifelong process: The German National Educational Panel Study (NEPS) [Special Issue]: Zeitschrift für Erziehungswissenschaft* (pp. 51–65, Vol. 14). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/s11618-011-0181-8

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.

Deville, J.-C., & Särndal, C.-E. (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association*, *87*(418), 376–382.

Haziza, D., & Beaumont, J.-F. (2017). Construction of weights in surveys: A review. *Statistical Science*, *32*(2), 206–226. https://doi.org/10.1214/16-sts608

Kalton, G., & Flores-Cervantes, I. (2003). Weighting methods. *Journal of Official Statistics*, *19*(2), 81–97.

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

Kultusministerkonferenz. (2012). Definitionenkatalog zur Schulstatistik 2012. http://www.kmk.org/fileadmin/pdf/Statistik/Defkat%5Ctextunderscore%202012.2%5Ctextunderscore%20m%5Ctextunderscore%20Anlagen.pdf

Särndal, C.-E., Swensson, B., & Wretman, J. (1992). *Model assisted survey sampling*. Springer.

Steinhauer, H. W., Aßmann, C., Zinn, S., Goßmann, S., & Rässler, S. (2015). Sampling and Weighting Cohort Samples in Institutional Contexts. *AStA Wirtschafts- und Sozialstatistisches Archiv*, *9*(2), 131–157. https://doi.org/10.1007/s11943-015-0162-0

Steinhauer, H. W., & Zinn, S. (2016). *NEPS Technical Report for Weighting: Weighting the Sample of Starting Cohort 3 of the National Educational Panel Study (Waves 1 to 5)* (Technical Report). Leibniz Institute for Educational Trajectories. Bamberg. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC3/5-0-0/SC3_5-0-0_W.pdf

Valliant, R., Dever, J. A., & Kreuter, F. (2013). *Practical tools for designing and weighting survey samples*. Springer.

Appendix

8.1 Tables

Table 1: Variables included in the weighting data set for SC8 SUF version 1.0.0 of the SUF

Variable	Applies to	Content
ID_t	6141	Identifier for target person
ID_i	6141	Identifier for the school the target person was initially sampled in
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_cal	6141	Final weight for target, calibrated
w_t_cal_std	6141	Final weight for target, calibrated and standardized
Final weights for wave-specific participants, calibrated and standardized		
w_t1	5763	Cross-sectional weight for targets participating in Wave 1

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	–	n.a.
	A104	6141	0	6141	5666	475	–	n.a.
	L006	475	8	467	97	370	–	n.a.

n.a. - “not (yet) available”.

Table 3: School models for panel entry.

	FS	GY	HS	IG	MB	OS	RS
(Intercept)	−2.318***	−1.065***	−1.886***	−1.499***	−1.467***	−0.965	−1.637***
	(0.414)	(0.286)	(0.385)	(0.345)	(0.420)	(0.640)	(0.483)
Classes in Grade 5	−0.251	−0.152	0.020	0.007	0.093	0.057	0.132
	(0.143)	(0.163)	(0.195)	(0.082)	(0.175)	(0.108)	(0.274)
Students in Grade 5	0.030	0.004	−0.004	0.001	−0.011	−0.016*	−0.009
	(0.016)	(0.006)	(0.010)	(0.003)	(0.008)	(0.008)	(0.010)
Funding: private (Ref. public)	0.233	−0.236	0.229	−0.318	0.079	−0.811	−0.228
	(0.316)	(0.188)	(0.243)	(0.256)	(0.252)	(0.518)	(0.340)
Regional classification: rural (Ref. urban)	0.778*	−0.046	−0.154	−0.148	0.507	−0.237	0.362
	(0.322)	(0.244)	(0.243)	(0.256)	(0.290)	(0.413)	(0.329)
Regional classification: sub-urban (Ref. urban)	0.085	−0.110	−0.025	−0.128	0.437	−0.033	0.182
	(0.220)	(0.124)	(0.196)	(0.159)	(0.277)	(0.336)	(0.246)
Sample position: Replacement 1 (Ref. Original)	0.733*	0.135	0.330	−0.324	−0.015	−0.013	0.149
	(0.310)	(0.247)	(0.378)	(0.332)	(0.329)	(0.615)	(0.414)
Sample position: Replacement 2 (Ref. Original)	0.331	−0.261	0.026	−0.196	−0.024	−0.078	−0.259
	(0.337)	(0.280)	(0.418)	(0.318)	(0.330)	(0.624)	(0.497)
Sample position: Replacement 3 (Ref. Original)	0.303	0.084	0.025	0.303	−0.263	−0.036	0.030
	(0.336)	(0.251)	(0.418)	(0.277)	(0.360)	(0.632)	(0.437)
Sample position: Replacement 4-9 (Ref. Original)	0.376	0.025	0.275	−0.098	−0.249	0.145	−0.123
	(0.348)	(0.196)	(0.313)	(0.231)	(0.262)	(0.461)	(0.343)
Log-likelihood	−107.871	−295.047	−136.012	−186.186	−145.002	−51.419	−83.269
N	590	1010	902	885	701	320	660

Significance: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$.

Table 4: Distributions in Grade 5 at student level.

	Frame		Gross		Panel		Part. wave 1	
	abs.	**%**	**abs.**	**%**	**abs.**	**%**	**abs.**	**%**
BW	102429	13.5	1504	7.3	464	7.6	441	7.7
BB	23559	3.1	361	1.8	149	2.4	146	2.5
BY	121327	16	2835	13.8	750	12.2	710	12.3
BE	33316	4.4	596	2.9	220	3.6	211	3.7
HB	6000	0.8	131	0.6	15	0.2	13	0.2
HH	16465	2.2	1464	7.1	226	3.7	208	3.6
HE	55941	7.4	1311	6.4	305	5	291	5
MV	14595	1.9	214	1	50	0.8	43	0.7
NI	72189	9.5	2437	11.9	703	11.4	658	11.4
NW	164391	21.7	3781	18.4	1347	21.9	1274	22.1
RP	36290	4.8	1023	5	298	4.9	268	4.7
SL	8274	1.1	129	0.6	5	0.1	4	0.1
SN	38228	5	1097	5.3	252	4.1	244	4.2
ST	19025	2.5	1118	5.4	500	8.1	464	8.1
SH	26061	3.4	1692	8.2	472	7.7	442	7.7
TH	19251	2.5	842	4.1	385	6.3	346	6
School branch								
GY	290312	38.3	9701	47.2	3090	50.3	2918	50.6
HS	51152	6.8	1041	5.1	377	6.1	351	6.1
MB	84223	11.1	2092	10.2	658	10.7	617	10.7
RS	114328	15.1	1503	7.3	462	7.5	440	7.6
IG	143210	18.9	5013	24.4	1089	17.7	1006	17.5
FS	16711	2.2	565	2.8	241	3.9	213	3.7
OS	57405	7.6	620	3	224	3.6	218	3.8
Gender								
Female	365802	48.3	9947	48.4	3046	49.6	2862	49.7
Male	391539	51.7	10406	50.7	3094	50.4	2900	50.3
Missing	0	0	182	0.9	1	0	1	0
Sum	**757341**	**100**	**20535**	**100**	**6141**	**100**	**5763**	**100**

Sources: School frame 2022/2023, A104_I_Bruttodaten_BY_SN, A104_I_Bruttodaten_HE, A104_T_TR.

Table 5: Summary statistics for all weights provided.

Label of weight	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
w_i	0.6148	1.8390	2.1139	2.4196	2.6423	5.2847	
w_t_cal	24.5004	76.9647	110.1265	123.3254	145.5434	416.3668	
w_t_cal_std	0.1987	0.6241	0.8930	1.0000	1.1802	3.3762	
w_t1	0.2076	0.6150	0.8856	1.0000	1.1635	3.1985	378

Table 6: Model estimating the individual participation propensities for being part of the panel sample, and model estimating the individual participation propensities for students in wave 1 of SC8 (excluding BY, SN, HE).

	Panel	Wave 1
Constant	−0.700***	2.710***
	(0.139)	(0.256)
School type: HS (ref. = "GY")	0.813**	−0.209
	(0.344)	(0.458)
School type: MB (ref. = "GY")	0.058	−0.098
	(0.206)	(0.264)
School type: RS (ref. = "GY")	0.065	−0.082
	(0.323)	(0.403)
School type: IG/FW (ref. = "GY")	−0.522***	−0.283
	(0.180)	(0.220)
School type: OS (ref. = "GY")	0.038	0.822*
	(0.290)	(0.494)
School type: FS (ref. = "GY")	1.213***	−0.376
	(0.279)	(0.446)
Age group: 11 years + (ref. = "10 years -")	−0.119***	−0.178
	(0.042)	(0.136)
Gender: Male (ref. = "Female")	−0.050	0.035
	(0.038)	(0.121)
Nationality: Non-German (ref. = "German")	−0.376***	−0.209
	(0.071)	(0.209)
Nationality: Missing (ref. = "German")	0.117	−0.738*
	(0.231)	(0.393)
Special educational needs: Yes (ref. = "No")	−0.121	−0.164
	(0.108)	(0.309)
Special educational needs: Missing (ref. = "No")	0.420	0.903
	(0.357)	(0.865)
Partial performance disorder: Yes (ref. = "No")	0.234**	−0.155
	(0.099)	(0.272)
Partial performance disorder: Missing (ref. = "No")	−0.645*	0.621
	(0.365)	(0.716)
School enrolment: Regular (ref. = "Earlier")	0.226***	0.336*
	(0.072)	(0.196)
School enrolment: Later (ref. = "Earlier")	0.189	−0.269
	(0.241)	(0.597)
School enrolment: Missing (ref. = "Earlier")	−0.768***	0.756*
	(0.169)	(0.437)
Observations	15,292	4,834

Notes: *p<0.1; **p<0.05; ***p<0.01; standard errors are given in parentheses. Source: A104_T_TR.