

NEPS

National Educational Panel Study

Research Data

Samples, Weights and Nonresponse: the Kindergarten Cohort of the National Educational Panel Study (Waves 1 to 5)

Würbach, A., Steinhauer, H.W., Zinn. S.

LifBi

LEIBNIZ INSTITUTE FOR
EDUCATIONAL TRAJECTORIES

Copyrighted Material
Leibniz Institute for Educational Trajectories (LifBi)
Wilhelmsplatz 3, 96047 Bamberg
Director: Prof. Dr. Hans-Günther Roßbach
Executive Director of Research: Dr. Jutta von Maurice
Executive Director of Administration: Dr. Robert Polgar
Bamberg, 2017

Samples, Weights, and Nonresponse: the Kindergarten Cohort of the National Educational Panel Study (Waves 1 to 5)

*Würbach, A., Steinhauer, H. W., & Zinn, S.
Leibniz Institute for Educational Trajectories*

Technical Report referring to [DOI:10.5157/NEPS:SC2:5.0.0](https://doi.org/10.5157/NEPS:SC2:5.0.0)

E-mail address of lead author:

methoden@lifbi.de

Samples, Weights, and Nonresponse: the Kindergarten Cohort of the National Educational Panel Study (Waves 1 to 5)

1 Prequel

The National Educational Panel Study (NEPS) surveys a cohort sample of Kindergarten children and Grade 1 students (Starting Cohort 2, SC2) and follows them over their transition to elementary school and beyond. The data are released via corresponding Scientific Use Files (SUF). The current SUF version is available under [DOI:10.5157/NEPS:SC2:5.0.0](https://doi.org/10.5157/NEPS:SC2:5.0.0).¹

This paper supplements the previous NEPS Working Paper by Steinhauer, Zinn, Gaasch, and Goßmann (2016) as well as the Technical Report by Steinhauer and Zinn (2016), which give detailed information on the applied indirect sampling procedure, the derivation of design weights, their successive adjustments, and the derivation of panel weights for previous waves.

In 2013, the cohort of Kindergarten children transitioned to elementary school. Children who transitioned to previously sampled schools were followed up within their institutional context together with their classmates who augment the cohort sample. Besides that, there are previously sampled schools no children transitioned to. Students within these schools also augment the cohort sample. Children who transitioned to other schools were tracked individually. By design, these children do not take part in the tests until Wave 6, when most of the children will be in Grade 4. Here, the entire sample will be surveyed and tested again. Table 1 documents the accordant study numbers and survey year available in the current SUF.

Along the distinct panel waves, for all participating children cross-sectional and longitudinal weights are provided. Cross-sectional weights are assigned to children relying on their participation in the different panel waves. Furthermore, weights are given for subgroups of the panel cohorts that are of special interest in our analysis. This concerns particularly the group of children continually taking part in the successive waves of the survey (currently, Waves 1 to 5) and the group of children and parents participating jointly. Longitudinal weights are provided for those children who have continually participated. Additional cross-sectional weights are provided for joint participation of children and parents (currently for Waves 1 to 3).

The remainder of this supplement is structured as follows: Section 2 details the panel progress as well as the new features of the corresponding weighting data sets. In Section 3 the non-response in Wave 5 is analyzed. Nonresponse models are estimated using probit regressions. Finally, Section 4 concludes with a summary of the provided sampling weights and design information given in the corresponding weighting data sets.

¹For general information on the NEPS, see Blossfeld, Roßbach, and von Maurice (2011). More detailed information is available in the documentation section on the [homepage](#).

Table 1: Survey overview for Starting Cohort 2.

Wave	Time	Study number
<i>Kindergarten children</i>		
1	2011	A12
2	2012	A13
<i>Elementary school students</i>		
3	2013	A14, A14A
4	2013	A15
5	2014	A89

2 Panel progress

The following Table 2 details the panel progress of Starting Cohort 2 by differentiating participants, temporary dropouts, and final dropouts. Final dropouts are separated into final dropouts due to refusal during the survey period and final dropouts between two consecutive waves.

Table 2: Panel progress of Starting Cohort 2 by wave.

Wave	Group	Panel Cohort			Status at the end of the wave			
		Total size	Not used	Used sample	Participants	Temporary dropout	Final dropout (in wave)	Final dropout (after wave)
1	All	^a 3007	0	3007	2949	47	11	0
2	All	2996	^b 215	2781	2727	54	0	1
3	All	9336	2419	6917	6733	184	0	5
	1	6341	0	6341	6176	165	0	2
	2	2419	2419	-	-	-	-	^c 3
4	3	576	0	576	557	19	0	0
	All	9331	2733	6598	^d 6340	232	26	22
	1	6339	296	6043	5801	217	25	15
	2	2416	2416	-	-	-	-	1
5	3	576	21	555	539	15	1	6
	All	9283	3119	6164	5799	204	161	79 ^e
	1	6299	669	5630	5296	185	149	42 ^e
	2	2415	2415	-	-	-	-	32 ^e
	3	569	35	534	503	19	12	5 ^e

Notes: "-" does not apply. Group 1 - The group of students tested in Grade 1 in elementary schools, who were not tested in Kindergarten institutions in Wave 1 and Wave 2. These (target) persons form the augmentation sample of Wave 3. Group 2 - The group of Kindergarten children who were tested only in Kindergartens in Wave 1 and Wave 2. In Wave 3, they are assigned to the individual retracking field and are temporary dropouts by design until Wave 6. Group 3 - The group of Kindergarten children, who were tested in Kindergartens in Wave 1 and Wave 2 and transition to elementary schools surveyed by NEPS in Wave 3. These (target) persons belong to the longitudinal sample of Waves 1, 2, and 3. ^a Panel size in Wave 1 is larger than the number of cases in the SUF, because of 11 final dropouts after Wave 1 and before publication of the SUF. Thus these cases are not included. ^b Cases not used left the institution they were surveyed in. These cases are tracked individually and surveyed again in Wave 6, when they are supposed to be in Grade 4. In the SUF their status is temporary dropout. ^c Final dropout in Group 2 is not included in the SUF. Here these cases are labeled as temporary dropout. ^d This table contains two cases less than in tx80220. These two cases were erroneously labeled as participants in the SUF SC2_5-0-0 (please compare the accordant release notes). ^e These numbers of final dropouts might change due to final assessment of the used samples in Wave 6 and Wave 7.

Compared to previous releases of the SC2 SUF the current weighting data sets contain an additional calibration of Group 1 in Wave 3 and an additional design variable:

The (target) persons of Group 1 form the augmentation sample in Wave 3. Data from Official Statistics (Statistisches Bundesamt, Fachserie 11, Reihe 1, 2012/13) regarding the gender ratio in the federal states has been used for post-stratification. By now each of the subgroups has panel entry weights calibrated to population size. The resulting (calibrated) weight `w_t_cal` is used as the basis for all succeeding wave-specific adjustments. Thus, derived cross-sectional and longitudinal weights are also calibrated.

The weighting data set contains an additional design variable `tx80112_R` which supplements the variable `h227108` from surveyed data in the module `pInstitution`. Both variables provide information regarding the total amount of students in each Grade 2 in the sample.

3 Participation in Wave 5

The processing in the nonresponse analysis is detailed in Chapter 3 in Steinhauer et al. (2016) as well as in Steinhauer and Zinn (2016). That is, a multilevel probit model is used to estimate the individual participation propensities for students in Grade 3 (participants in Wave 5) separately for the Groups 1 and 3. The results are given in Table 3. As can be seen, participation in previous waves significantly influences the participation probability in the current wave in Group 1.

Please refer to Steinhauer et al. (2016) and Steinhauer and Zinn (2016) regarding the interpretation of the participation propensity in previous waves.

On the basis of the estimated (non)response models participation probabilities are predicted and used as adjustment factors to derive cross-sectional and longitudinal survey weights.

4 Summary of Weights

The NEPS provides various kinds of weights for Kindergarten children and elementary school students together with design information. Table 4 lists the design information and the different weights provided by SUF release version [DOI:10.5157/NEPS:SC2:5.0.0](https://doi.org/10.5157/NEPS:SC2:5.0.0). In SC2, weights are provided in two distinct weighting files: one for Kindergarten children (Groups 2 and 3) and one for elementary school students (Groups 1 and 3). All cross-sectional and longitudinal weights are provided in a trimmed and standardized form. The trimmed sampling weights are standardized with mean 1 to ease statistical weighted analysis (cp. Chapter 4 in Steinhauer and Zinn (2016)). Summary statistics for all kind of weights provided are given in Table 5.

Please refer to Chapter 6 in Steinhauer and Zinn (2016) for advices regarding the usage of weights.

Table 3: Models estimating the individual participation propensities for Kindergarten children in Wave 1, Wave 2, and students in Grade 1 in Wave 3, Wave 4 and Wave 5 of SC2 used to derive adjustment factors for adjusted wave-specific cross-sectional and longitudinal weights.

	Wave 1		Wave 2		Wave 3		Wave 4		Wave 5	
	Group 1	Group 3	Group 1	Group 3	Group 1	Group 3	Group 1	Group 3	Group 1	Group 3
(Intercept)	3.472*** (0.295)	3.061*** (0.380)	1.839*** (0.101)	0.014 (0.279)	1.898*** (0.108)	1.701*** (0.182)	1.571*** (0.087)	0.062 (0.199)		
Place of residence with both parents	0.317 (0.173)									
Participation in Wave 1		0.935** (0.330)								
Participation in Wave 3							0.421* (0.177)		0.384* (0.161)	
Participation in Wave 4									1.296*** (0.129)	
Native language German				1.258*** (0.146)						
Native language other than German				1.389*** (0.181)						
Special educational needs				1.502*** (0.277)						
Special educational needs				1.356*** (0.350)						
Random intercept (SD) on the Kindergarten level	1.953	2.468								
on the school level					0.794		0.623			0.370
Sample size	2996	2781	576	6341	555	6043	534	5630		

Notes: Reference categories are: Place of residence (not with both parents), Participation in Wave 1 (no), Native language (missing), Special educational needs (unknown). To model individual participation, the `glmex` function with a probit link provided by `lme4` package (Bates, Maechler, & Bolker, 2012) in R (R Core Team, 2015) was used. ***, **, and * denote significance at the 0.1%, 1%, and 5% level, respectively. Standard errors are given in parentheses.

Group 1 - The group of students tested in Grade 1 in elementary schools, who were not tested in Kindergarten institutions in Wave 1 and Wave 2. These (target) persons form the augmentation sample of Wave 3.

Group 3 - The group of Kindergarten children, who were tested in Kindergartens in Wave 1 and Wave 2 and transition to elementary schools surveyed by NEPS in Wave 3.

Table 4: Variables included in the weighting data for SC2 version 5.0.0 of the SUF.

Variable	Applies to	Content
<i>Identifier</i>		
ID_t	all targets	Identifier for target person
ID_i	all targets	Identifier for the school the target person was initially sampled in
<i>Design information</i>		
tstud_st	all targets	Study number the target person was first surveyed in (A12, A14, A14A)
group	all targets	Grouping variable for children in Kindergarten and school context
stratum_imp2_R	schools	Implicit sampling stratum (Federal State the school is located in according to sampling frame)
stratum_imp3_R	schools	Implicit sampling stratum (regional classification according to sampling frame)
stratum_imp4_R	schools	Implicit sampling stratum (funding according to sampling frame)
tx80112_R	schools	Total amount of students in Grade 2 (from Official Statistics)
<i>Weights referring to Kindergarten children (Groups 2 and 3)</i>		
w_i	2,996 cases	Nonresponse adjusted design weight for Kindergarten
w_t	2,996 cases	Nonresponse adjusted design weight for target
w_t_cal	2,996 cases	Calibrated nonresponse adjusted design weight for target
w_t1	2,949 cases	Cross-sectional weight for targets participating in Wave 1
w_tp1	2,309 cases	Cross-sectional weight for targets jointly participating with one parent in Wave 1
w_t2	2,727 cases	Cross-sectional weight for targets participating in Wave 2
w_tp2	1,965 cases	Cross-sectional weight for targets jointly participating with one parent in Wave 2
w_t12	2,685 cases	Longitudinal weight for targets participating in Wave 1 and 2
w_tp12	1,804 cases	Longitudinal weight for targets jointly participating with one parent in Wave 1 and 2
w_t123	539 cases	Longitudinal weight for targets participating in Wave 1, 2, and 3
w_tp123	388 cases	Longitudinal weight for targets jointly participating with one parent in Wave 1, 2, and 3
w_t1234	504 cases	Longitudinal weight for targets participating in Wave 1 up to Wave 4
w_t12345	460 cases	Longitudinal weight for targets participating in Wave 1 up to Wave 5
<i>Weights referring to elementary schools students (Groups 1 and 3)</i>		
w_i	6,917 cases	Nonresponse adjusted design weight for elementary school
w_t	6,917 cases	Nonresponse adjusted design weight for target (Grade 1 student)
w_t_cal	6,917 cases	Calibrated nonresponse adjusted design weight for target
w_t3	6,733 cases	Cross-sectional weight for targets participating in Wave 3
w_tp3	5,636 cases	Cross-sectional weight for targets jointly participating with one parent in Wave 3
w_t4	6,340 cases	Cross-sectional weight for targets participating in Wave 4
w_t5	5,799 cases	Cross-sectional weight for targets participating in Wave 5
w_t34	6,185 cases	Longitudinal weight for targets participating in Wave 3 and 4
w_t345	5,565 cases	Longitudinal weight for targets participating in Wave 3, 4 and 5

Table 5: Summary statistics for all weights provided.

Label of weight	Min.	Lower Quart.	Median	Mean	Upper Quart.	Max.
<i>Weights referring to Kindergarten children (Groups 2 and 3)</i>						
w_i	14.5722	76.0647	106.4483	138.2196	157.1446	1564.0796
w_t	21.4348	104.5298	156.6036	220.5345	246.0522	9644.9638
w_t_cal	9.1204	94.5444	143.5075	216.1175	248.5428	3269.7032
w_t1	0.0438	0.4558	0.6941	1.0000	1.2094	4.2676
w_tp1	0.0456	0.4594	0.7017	1.0000	1.2163	4.2532
w_t2	0.0412	0.4307	0.6777	1.0000	1.1863	4.4543
w_tp2	0.0297	0.3262	0.5155	1.0000	0.9454	5.2035
w_t12	0.0437	0.4549	0.7075	1.0000	1.2123	4.2548
w_tp12	0.0441	0.4733	0.7007	1.0000	1.1840	4.2237
w_t123	0.1129	0.4688	0.7410	1.0000	1.2311	4.1583
w_tp123	0.1139	0.4721	0.7385	1.0000	1.2160	4.0484
w_t1234	0.1122	0.4645	0.7309	1.0000	1.2427	4.2174
w_t12345	0.1112	0.4580	0.7310	1.0000	1.2312	4.1959
<i>Weights referring to elementary schools students (Groups 1 and 3)</i>						
w_i	9.4518	23.5050	30.9038	39.3148	43.3969	332.6401
w_t	15.4801	44.8275	67.9257	90.7713	106.4174	4000.4377
w_t_cal	16.6909	48.0933	73.3117	97.5463	115.7840	3869.2945
w_t3	0.1746	0.5063	0.7827	1.0000	1.2129	3.7875
w_tp3	0.1602	0.4938	0.7505	1.0000	1.2314	3.9137
w_t4	0.1591	0.4770	0.7316	1.0000	1.1715	4.1370
w_t34	0.1741	0.5131	0.7856	1.0000	1.2254	3.7025
w_t5	0.1471	0.4553	0.6887	1.0000	1.1581	4.3740
w_t345	0.1707	0.5183	0.7832	1.0000	1.2347	3.6993

Acknowledgements This paper uses data from the National Educational Panel Study (NEPS): Starting Cohort Kindergarten, DOI:10.5157/NEPS:SC2:5.0.0. From 2008 to 2013, NEPS data was collected as part of the Framework Program for the Promotion of Empirical Educational Research funded by the German Federal Ministry of Education and Research (BMBF). As of 2014, NEPS is carried out by the Leibniz Institute for Educational Trajectories (LifBi) at the University of Bamberg in cooperation with a nationwide network.

References

- Bates, D., Maechler, M., & Bolker, B. (2012). *lme4: Linear mixed-effects models using Eigen and Eigenfaces*. Retrieved 03.06.2014, from <http://CRAN.R-project.org/package=lme4>
- Blossfeld, H.-P., Roßbach, H. G., & von Maurice, J. (Eds.). (2011). *Education as a lifelong process: The German National Educational Panel Study (NEPS) [Special Issue]: Zeitschrift für Erziehungswissenschaft* (Vol. 14). Wiesbaden: VS Verlag für Sozialwissenschaften.
- R Core Team. (2015). *R: A language and environment for statistical computing* [Computer software manual]. Vienna, Austria. Retrieved from <https://www.R-project.org/>
- Statistisches Bundesamt. (2013). *Bildung und Kultur - Allgemeinbildende Schulen: Schuljahr 2012/2013* (Fachserie 11 No. Reihe 1). Wiesbaden: Statistisches Bundesamt. Retrieved 15.03.2017, from <https://www.destatis.de/DE/Publikationen/Thematisch/BildungForschungKultur/Schulen/AllgemeinbildendeSchulen2110100137004.pdf>
- Steinhauer, H. W., & Zinn, S. (2016). *NEPS Technical Report for Weighting: Weighting the sample of Kindergarten Children and Grade 1 Students of the National Educational Panel Study (Waves 1 to 4)* (Technical Report). Bamberg: Leibniz Institute for Educational Trajectories. Retrieved from https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC2/4-0-0/SC2_4-0-0_W.pdf
- Steinhauer, H. W., Zinn, S., Gaasch, C., & Goßmann, S. (2016). *NEPS Technical Report for Weighting: Weighting the sample of Kindergarten Children and Grade 1 Students of the National Educational Panel Study (Wave 1 to 3)* (Working Paper No. 66). Bamberg: Leibniz Institute for Educational Trajectories. Retrieved from https://www.neps-data.de/Portals/0/Working%20Papers/WP_LXVI.pdf