NEPS
National Educational Panel Study

Research Data

Würbach, A.

Samples, Weights and Nonresponse

NEPS Starting Cohort 2 — Kindergarten

From Kindergarten to Elementary School

Wave 7

LIfBi
LEIBNIZ INSTITUTE FOR
EDUCATIONAL TRAJECTORIES

Samples, Weights, and Nonresponse: the Kindergarten Cohort of the National Educational Panel Study (Wave 7)

Würbach, A.

Leibniz Institute for Educational Trajectories

Technical Report referring to DOI:10.5157/NEPS:SC2:7.0.0

E-mail address of lead author:

methoden@lifbi.de

Samples, Weights, and Nonresponse: the Kindergarten Cohort of the National Educational Panel Study (Wave 7)

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:SC2:7.0.0.[1]

This paper supplements the previous reports for weighting by Würbach (2018), Würbach, Steinhauer, and Zinn (2017) as well as the more detailed NEPS Working Paper by Steinhauer, Zinn, Gaasch, and Goßmann (2016) and the Technical Report by Steinhauer and Zinn (2016), which give 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 did not take part in the tests until Wave 6. Then, in Wave 6 (Grade 4), the entire sample was surveyed and tested again. As of Wave 7, surveying and testing was no longer done in the institutional context but in the individual retracking field for all targets.

Due to its composition the panel cohort of SC2 can be categorized into three groups:

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.

Table 1 documents the accordant study numbers and survey year available in the current SUF.

[1] For 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	Year	Time	Study number
Kindergarten children			
1	2011	4-5 years	A12
2	2012	5-6 years	A13
Elementary school students			
3	2013	Grade 1	A14, A14A
4	2013	Grade 2	A15, A15-L1
5	2014	Grade 3	A89
6	2015	Grade 4	A97, B103
7	2016	Grade 5	B104

For all participating children cross-sectional and, where appropriate, 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 and the group of children and parents participating jointly. Longitudinal weights are provided for those children who have continually participated. Additional cross-sectional and longitudinal weights are provided for joint participation of children and parents.

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. The weighting adjustments are described in Section 3. For this, a new Subsection 3.1 is introduced that depicts the models for describing the propensity to be a panel member at start of Wave 6. Probit regressions are used to estimate the individual panel entrance propensities. In the following Subsection 3.2 the nonresponse in Wave 7 and the response propensity for participating in consecutive waves is analyzed. Nonresponse models are estimated using probit regressions, too. Finally, Section 4 concludes with a summary of the provided sampling weights and design information given in the corresponding weighting data sets.

2 Panel progress

The following Table 2 details the panel progress of Starting Cohort 2 by differentiating participants, temporary dropouts, and final dropouts for each group separately and in total. 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.

		Panel Cohort			Status at the end of the wave			
Wave	Group	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
	3	576	0	576	557	19	0	0
4	**All**	9331	2733	6598	6340	232	26	23
	1	6339	296	6043	5801	217	25	15
	2	2416	2416	-	-	-	-	2
	3	576	21	555	539	15	1	6
5	**All**	9282	3118	6164	5799	204	161	77
	1	6299	669	5630	5296	185	149	41
	2	2414	2414	-	-	-	-	31
	3	569	35	534	503	19	12	5
6	**ALL**	9044	555	8489	6942	1180	367	[d]693
	1	6109	62	6047	5461	425	161	[d]185
	2	2383	458	1925	998	735	192	[d]497
	3	552	35	517	483	20	14	[d]11
7	**ALL**	7984	[e]34	7950	4220	3687	43	42
	1	5763	1	5762	3246	2483	33	30
	2	1694	0	1694	648	1036	10	9
	3	527	33	494	326	168	0	3

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] Among these final dropouts also previous parent withdrawals are subsumed. Due to the movement into the individual field the related students cannot be asked for participation again. [e] 34 target students could not have been surveyed and tested for technical reasons.

Compared to the previous release of the SC2 SUF (version 6.0.1), the current weighting data sets contain an additional weighting data set which starts in Wave 6 and is based on a composite design weight for all targets together again. This composite Grade 4 design weight is adjusted to nonresponse and calibrated to population size in Wave 6. Data from Official Statistics (Statistisches Bundesamt, Fachserie 11, Reihe 1, 2015/16) regarding the gender ratio in the federal states has been used for post-stratification. In Table 3 the corresponding figures are given. More information regarding the construction of composite design weights for different subsamples can be found in Zinn, Würbach, Steinhauer, and Kiesl (2018). Note that due to small sample sizes the former weighting data set containing the longitudinal weights for Kindergarten children only is frozen and not updated any further. In the data set for joint weights as well as in the weighting data set referring to elementary schools students (Groups 1 & 3) only standardized weights are given.

Table 3: Population sizes used for calibration of the Grade 4 students in Wave 6.

	Gender	
`stratum_imp2_R`	Female	Male
Brandenburg	9640	10268
Berlin	13624	14264
Baden-Württemberg	47500	49705
Bayern	53262	56066
Bremen	2621	2843
Hessen	26124	27435
Hamburg	7344	7682
Mecklenburg-Vorpommern	6141	6474
Niedersachsen	35937	38275
Nordrhein-Westfalen	77376	81468
Rheinland-Pfalz	16372	17629
Schleswig-Holstein	12029	12799
Saarland	3561	3871
Sachsen	15766	16391
Sachsen-Anhalt	8026	8417
Thüringen	8044	8442
Sum	343,367	362,029

3 Weighting Adjustments for Wave Participation

Systematic refusals may arise and for this, the (non)response and attrition processes of the sampled individuals, has to be accounted for. Thus, for reasons of usability, commonly design weights are adjusted to account for nonresponse in the survey. For this purpose, the units' probabilities to participate in each survey wave as well as in consecutive waves are employed. The processing in the nonresponse analysis is detailed in Chapter 3 in Steinhauer et al. (2016) as well as in Steinhauer and Zinn (2016). The following estimated (non)response models are used as basis for calculation of participation probabilities and hence serve as adjustment factors to derive cross-sectional and longitudinal survey weights.

3.1 Modeling Panel Participation in Wave 6

Probit models are used to estimate the individual panel entrance propensities for Wave 6. Directly on the onset of Wave 6, the panel cohort comprised 9,044 targets (6,109 in Group 1, 2,383 in Group 2, and 552 in Group 3). That is, meanwhile 293 panel members (or their parents) withdrew their participation consent either during fieldwork or between two consecutive waves. Table 4 gives the corresponding variables and results for panel participation separately for Groups 1 & 3 being tracked institutionally and Group 2 being tracked individually. Please note that the panel entry model for Groups 1 & 3 carry only these information being available for both subsamples together. The overlapping variables are gender (reference group: male), age group (reference group: older half), native language (German, other than German, reference group: missing), start schooling (earlier, regular, later, reference group: missing) and special educational needs (no, yes, reference group: unknown) with additional information regarding previous wave participation. The panel entry model for Group 2 comprises also gender and age group, but German spoken at home (reference group: no) and place of residence (reference group: not with both parents) as well as participation in Wave 1 and Wave 2.

Table 4: Models estimating the individual panel entrance propensities for respondents of the SC2 in Wave 6.

Value	**Group 1 & 3**	**Group 2**
(Intercept)	1.189^{***}	1.453^{***}
	(0.075)	(0.293)
Participation in Wave 1		0.737^{*}
yes		(0.300)
Participation in Wave 4	-0.542^{***}	
yes	(0.092)	
Participation in Wave 5	1.869^{***}	
yes	(0.080)	
Number of cases	6917	2420

Notes: Reference categories are: Participation in Wave 1/4/5 (no). ***, **, and * denote significance at the 0.1%, 1%, and 5% level, respectively. Standard errors are given in parentheses.

For Groups 1 & 3 only previous wave participation remains among the significant variables. Being participant in the immediately preceding wave shows a positive and highly significant effect on panel willingness. Participation in the wave before shows a significant effects too but pointing to the opposite direction. That is, participation in Wave 4 seems to decrease the probability to stay in the panel. However, we suggest that this effect is overpowered by participation in the immediately preceding wave. This can be supported by inspection of the Pseudo-R^2.[2] The model for panel entrance at Wave 6 for Groups 1 & 3 – as shown here – has a Pseudo-R^2 of 0.262. While the model containing only participation in Wave 5 has a Pseudo-R^2 of 0.260 indicating that the largest contribution to explain the panel entrance propensity stems from one variable only.

[2] For calculation of the Pseudo-R^2, the `r.squaredGLMM` function in the `MuMIn` package (Barton, 2017) in R (R Core Team, 2018) was used.

None of the respondent characteristics were significant at least at a 5% level for Group 2, too. Here, participation in the first wave shows a positive and significant effect on panel willingness at start of Wave 6. This is highly plausible because Group 2 was not surveyed in Wave 3 to 5.

3.2 Modeling Participation in Wave 7

To estimate the individual participation propensities for students in Grade 5 (participants in Wave 7) also a (multilevel) probit model is used.[3] The results are given in Table 5. As can be seen, participation in previous waves highly influences the participation probability in the current wave in Group 1 and Group 2. In Group 3 the probability of attending the CAPI (computer-assisted personal interview) is significantly influenced by the place of residence. Students living together with both parents are more likely to participate than students living with a single parent. This holds also for Group 2, but here the language spoken at home additionally influences the propensity to participate. If German is spoken at home the likelihood for participation is significantly increased. In Group 1 the language is of relevance as well. The language spoken at home is not available for this group but the native language is known instead. As can be seen, if German is the native language the student is more likely to participate than students with missing language information. On the opposite, students with a native language different from German are less likely to participate than students for which the language information is missing. Group 1 also exhibits an effect of age. The younger half of students is more willing to participate than the older half.

Joint participation propensities for targets and one parent are given in Table 6. In all three groups parent participation in previous waves is highly significant for joint participation. For Groups 1 and 2 also target participation in the immediately preceding wave shows a significant effect on the willingness to participate jointly in Wave 7. For Group 3 it is the information living together with both parents that is significant with respect to joint participation propensity but is less significant.

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

[3] To model individual participation in the school context, the `glmer` function with a probit link provided by `lme4` package (Bates, Maechler, Bolker, & Walker, 2015) in R (R Core Team, 2018) was used.

Table 5: Models estimating the individual participation propensities for students in Grade 5 (Wave 7) of SC2 used to derive adjustment factors for adjusted wave-specific cross-sectional and longitudinal weights.

		Wave 7	
Value	Group 3	Group 2	Group 1
(Intercept)	−0.057	−1.707***	−0.983***
	(0.189)	(0.129)	(0.095)
Target age group			0.134***
younger half			(0.035)
Place of residence	0.518**	0.351***	
with both parents	(0.199)	(0.092)	
Native language			0.191***
German			(0.057)
Native language			−0.219***
other than German			(0.066)
German spoken at home		0.469***	
yes		(0.098)	
Participation in Wave 5			0.192**
yes			(0.063)
Participation in Wave 6		1.125***	0.903***
yes		(0.072)	(0.080)
Random intercept (SD)			
on the school level			0.239
Sample size	494	1694	5762

Notes: Reference categories are: Target age group (older half), Place of residence (not with both parents), Native language (missing), German spoken at home (no), Participation in Wave 5/6 (no).
***, **, 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 but not being tested in Kindergarten institutions in Wave 1 and 2 (forming the augmentation sample of Wave 3).
Group 2 - The group of Kindergarten children individually tested in Wave 6.
Group 3 - The group of Kindergarten children being tested in Kindergartens in Wave 1 and Wave 2 and transition to elementary schools surveyed by NEPS in Wave 3.

Table 6: Models estimating the joint participation propensities for students and parents in Wave 7 of SC2 used to derive adjustment factors for adjusted wave-specific cross-sectional and longitudinal weights

	Group 3		Group 2		Group 1	
	Student	Parent	Student	Parent	Student	Parent
(Intercept)	-0.897^{***}	-2.555^{***}	-1.501^{***}	-2.400^{***}	-1.190^{***}	-2.442^{***}
	(0.225)	(0.342)	(0.085)	(0.147)	(0.077)	(0.121)
Place of residence	0.474^{*}	0.666^{*}				
With both parents	(0.214)	(0.259)				
German spoken at home				0.363^{**}		
yes				(0.130)		
Target part. in Wave 6			0.791^{***}	0.630^{***}	0.595^{***}	0.512^{***}
yes			(0.078)	(0.096)	(0.076)	(0.101)
Parent part. in Wave 2			0.433^{***}			
yes			(0.087)			
Parent part. in Wave 3		0.537^{*}				0.310^{***}
yes		(0.214)				(0.073)
Parent part. in Wave 4				0.525^{***}	0.187^{***}	0.443^{***}
yes				(0.103)	(0.052)	(0.066)
Parent part. in Wave 5	0.371^{*}	0.849^{***}		0.596^{***}	0.386^{***}	0.764^{***}
yes	(0.175)	(0.201)		(0.104)	(0.051)	(0.057)
Parent part. in Wave 6	0.955^{***}	1.766^{***}	0.787^{***}	1.502^{***}	0.597^{***}	1.553^{***}
yes	(0.171)	(0.198)	(0.076)	(0.098)	(0.049)	(0.054)
Correlation	0.383		0.325		0.307	
Sample size	494		1694		5762	

Notes: Reference categories are: Place of residence (not with both), German spoken at home (no), Target participating in Wave 6 (no), Parent participating in Wave (no). ***, **, and * denote significance at the 0.1%, 1%, and 5% level, respectively. Standard errors are given in parentheses.
For modeling joint participation decisions, the `SemiParBIVProbit` function in the same named package (Marra & Radice, 2013; Radice, Marra, & Wojtys, 2016) in R was used (R Core Team, 2018).

4 Summary of Weights

The NEPS provides various kinds of weights for Kindergarten children and elementary school students together with design information. Table 7 lists the design information and the different weights provided by SUF release version DOI:10.5157/NEPS:SC2:7.0.0. In SC2, weights are provided in three distinct weighting files: one for Kindergarten children (Groups 2 & 3), which is frozen in Wave 6 and will not be continued, one for elementary school students (Groups 1 & 3), and one for Grade 4 students transferring to lower secondary education (Sec I). The weighting data set referring to Kindergarten children (Groups 2 & 3) provides all cross-sectional and longitudinal weights in a trimmed and standardized form. Weights are standardized with mean 1 to ease statistical weighted analysis, cp. Chapter 4 in Steinhauer and Zinn (2016). The weighting data set referring to elementary schools students (Groups 1 & 3) and the data set with joint weights provide all cross-sectional and longitudinal weights standardized.

Table 7: Variables included in the weighting data sets for SC2 Version 7.0.0 of the SUF.

Variable	Applies to	Content
Identifier		
ID_t	all targets	Identifier for target person
ID_i	all targets	Identifier for the school the target person was initially sampled in
Design information		
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)
Weights referring to Kindergarten children (Groups 2 & 3)		
w_i	2,996 cases	Nonresponse adjusted design weight for Kindergarten, with $i = 1, \ldots, 268$
w_t	2,996 cases	Calibrated nonresponse adjusted design weight for target (Kindergarten child)
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_tp1234	335 cases	Longitudinal weight for targets jointly participating with one parent in Wave 1 up to Wave 4
w_t12345	460 cases	Longitudinal weight for targets participating in Wave 1 up to Wave 5
w_tp12345	276 cases	Longitudinal weight for targets jointly participating with one parent in Wave 1 up to Wave 5
w_t123456	433 cases	Longitudinal weight for targets participating in Wave 1 up to Wave 6
w_tp123456	238 cases	Longitudinal weight for targets jointly participating with one parent in Wave 1 up to Wave 6

Table 7: Variables included in the weighting data sets for SC2 Version 7.0.0 of the SUF (cont.).

Variable	Applies to	Content
Weights referring to elementary schools students (Groups 1 & 3)		
w_i	6,917 cases	Nonresponse adjusted design weight for elementary school, $i = 1, \ldots, 279$
w_t	6,917 cases	Calibrated nonresponse adjusted design weight for target (Grade 1 student)
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_tp4	4,865 cases	Cross-sectional weight for targets jointly participating with one parent in Wave 4
w_t34	6,189 cases	Longitudinal weight for targets participating in Wave 3 and 4
w_tp34	4,487 cases	Longitudinal weight for targets jointly participating with one parent in Wave 3 and 4
w_t5	5,799 cases	Cross-sectional weight for targets participating in Wave 5
w_tp5	4,025 cases	Cross-sectional weight for targets jointly participating with one parent in Wave 5
w_t345	5,567 cases	Longitudinal weight for targets participating in Wave 3, 4 and 5
w_tp345	3,500 cases	Longitudinal weight for targets jointly participating with one parent in Wave 3, 4 and 5
w_t6	6,942 cases	Cross-sectional weight for targets participating in Wave 6
w_tp6	4,641 cases	Cross-sectional weight for targets jointly participating with one parent in Wave 6
w_t3456	5,256 cases	Longitudinal weight for targets participating in Wave 3 up to Wave 6
w_tp3456	3,046 cases	Longitudinal weight for targets jointly participating with one parent in Wave 3 up to Wave 6
w_t7	4,220 cases	Cross-sectional weight for targets participating in Wave 7
w_tp7	3,247 cases	Cross-sectional weight for targets jointly participating with one parent in Wave 7
w_t34567	3,093 cases	Longitudinal weight for targets participating in Wave 3 up to Wave 7
w_tp34567	2,098 cases	Longitudinal weight for targets jointly participating with one parent in Wave 3 up to Wave 7
Weights referring to Grade 4 to Sec I students (Groups 1, 2 and 3)		
w_p6	9,044 cases	Calibrated panel entry weight for target (Grade 4 student)
w_p6_joint	9,044 cases	Calibrated joint panel entry weight for target (Grade 4 student)
w_t6	6,942 cases	Cross-sectional weight for targets participating in Wave 6
w_tp6	4,641 cases	Cross-sectional weight for targets jointly participating with one parent in Wave 6
w_t7	4,220 cases	Cross-sectional weight for targets participating in Wave 7
w_tp7	3,247 cases	Cross-sectional weight for targets jointly participating with one parent in Wave 7
w_t67	4,015 cases	Longitudinal weight for targets participating in Wave 6 and 7
w_tp67	2,988 cases	Longitudinal weight for targets jointly participating with one parent in Wave 6 and 7

Summary statistics for all kind of weights provided are given in Table 8.

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

Table 8: Summary statistics for all weights provided.

Label of weight	Min.	Lower Quart.	Median	Mean	Upper Quart.	Max.
Weights referring to Kindergarten children (Groups 2 & 3)						
w_i	14.5722	76.0647	106.4482	138.2196	157.1446	1564.0796
w_t	9.1204	94.5444	143.0656	216.1175	248.5428	3269.7032
w_t1	0.0438	0.4559	0.6929	1.0000	1.2097	4.2678
w_tp1	0.0456	0.4616	0.7029	1.0000	1.2175	4.2533
w_t2	0.0412	0.4308	0.6768	1.0000	1.1866	4.4548
w_tp2	0.0297	0.3254	0.5160	1.0000	0.9508	5.1988
w_t12	0.0437	0.4550	0.7055	1.0000	1.2126	4.2555
w_tp12	0.0442	0.4735	0.6982	1.0000	1.1812	4.2431
w_t123	0.1122	0.4660	0.7365	1.0000	1.2241	4.1100
w_tp123	0.1017	0.4696	0.7343	1.0000	1.2093	4.0561
w_t1234	0.1115	0.4622	0.7261	1.0000	1.2360	4.1736
w_tp1234	0.0983	0.4546	0.7259	1.0000	1.2449	4.1839
w_t12345	0.1105	0.4582	0.7269	1.0000	1.2255	4.1512
w_tp12345	0.0983	0.4539	0.7352	1.0000	1.2469	4.1273
w_t123456	0.1123	0.4666	0.7355	1.0000	1.2355	4.0926
w_tp123456	0.0994	0.4597	0.7189	1.0000	1.2603	4.1200
Weights referring to elementary schools students (Groups 1 & 3)						
w_i	9.4518	23.5052	30.9038	39.3148	43.3969	332.6401
w_t	16.6683	47.9776	73.3316	97.5463	115.7840	3869.2945
w_t3	0.1663	0.4855	0.7448	1.0000	1.1822	32.1602
w_tp3	0.1163	0.3595	0.5512	1.0000	0.8984	348.1264
w_t4	0.1077	0.3227	0.4941	1.0000	0.7991	146.3842
w_tp4	0.0275	0.0910	0.1407	1.0000	0.2564	1003.5271
w_t34	0.1685	0.4984	0.7605	1.0000	1.1971	13.7567
w_tp34	0.1609	0.4973	0.7396	1.0000	1.1907	18.8211
w_t5	0.0862	0.2657	0.4089	1.0000	0.6784	128.6669
w_tp5	0.0090	0.0311	0.0501	1.0000	0.1020	1591.6348
w_t345	0.1654	0.4976	0.7582	1.0000	1.1912	13.7615
w_tp345	0.1553	0.4804	0.7376	1.0000	1.1934	18.5344
w_t6	0.0525	0.1848	0.3004	1.0000	0.5810	184.3338
w_tp6	0.0016	0.0086	0.0149	1.0000	0.0420	1264.8013
w_t3456	0.1640	0.4964	0.7584	1.0000	1.1971	13.7397
w_tp3456	0.1531	0.4837	0.7292	1.0000	1.1820	18.2809
w_t7	0.0310	0.1178	0.2041	1.0000	0.4257	445.4226
w_tp7	0.0013	0.0052	0.0095	1.0000	0.0319	1361.7294
w_t34567	0.1482	0.4842	0.7566	1.0000	1.2311	16.6198
w_tp34567	0.1559	0.4844	0.7238	1.0000	1.2051	18.6151
Joint weights referring to Grade 4 students (Groups 1, 2 and 3)						
w_p6	11.8518	58.1023	95.3552	155.9920	173.8639	4501.0276
w_p6_joint	3.1234	37.0489	56.5586	77.9960	92.1824	2832.7848
w_t6	0.0643	0.4525	0.7082	1.0000	1.1708	19.6040
w_tp6	0.0373	0.3347	0.5227	1.0000	0.9312	106.7981
w_t7	0.0630	0.2710	0.4490	1.0000	0.7675	46.7226
w_tp7	0.0221	0.1242	0.2009	1.0000	0.3912	175.8683
w_t67	0.1019	0.4264	0.6815	1.0000	1.1325	26.8890
w_tp67	0.0611	0.3277	0.5099	1.0000	0.9091	73.0282

Acknowledgements This paper uses data from the National Educational Panel Study (NEPS): Starting Cohort Kindergarten, DOI:10.5157/NEPS:SC2:7.0.0. 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

Barton, K. (2017). MuMIn: Multi-Modal Inference [Computer software manual]. Retrieved from http://CRAN.R-project.org/package=MuMIn (R package version 1.40.0)

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). *Fitting Linear Mixed-Effects Models using lme4* (Vol. 67) (No. 1). Retrieved 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.

Marra, G., & Radice, R. (2013). A penalized likelihood estimation approach to semiparametric sample selection binary response modeling. *Electronic Journal of Statistics*, *7*, 1432-1455. Retrieved from https://www.ucl.ac.uk/statistics/research/pdfs/rr315.pdf

R Core Team. (2018). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from https://www.R-project.org/

Radice, R., Marra, G., & Wojtys, M. (2016). Copula regression spline models for binary outcomes. *Statistics and Computing*, *26*(5), 981–995. doi: 10.1007/s11222-015-9581-6

Statistisches Bundesamt. (2016). *Bildung und Kultur: Allgemeinbildende Schulen, Fachserie 11, Reihe 1* (Tech. Rep.). Wiesbaden: Destatis: Statistisches Bundesamt. Retrieved from https://www.destatis.de/DE/Publikationen/Thematisch/BildungForschungKultur/Schulen/AllgemeinbildendeSchulen2110100167004.pdf?__blob=publicationFile

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

Würbach, A. (2018). *Samples, Weights, and Nonresponse: the Kindergarten Cohort of the National Educational Panel Study (Waves 1 to 6)* (Technical Report). Bamberg: Leibniz Institute for Educational Trajectories. Retrieved from https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC2/6-0-1/SC2_6-0-1_W.pdf

Würbach, A., Steinhauer, H. W., & Zinn, S. (2017). *Samples, Weights, and Nonresponse: the Kindergarten Cohort of the National Educational Panel Study (Waves 1 to 5)* (Technical Report). Bamberg: Leibniz Institute for Educational Trajecto-

ries. Retrieved from https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC2/5-0-0/SC2_5-0-0_W.pdf

Zinn, S., Würbach, A., Steinhauer, H. W., & Kiesl, H. (2018). *The Composite Grade 4 Weight of the Kindergarten Cohort of the National Educational Panel Study* (NEPS Survey Paper No. 44). Bamberg: Leibniz Institute for Educational Trajectories, National Educational Panel Study. Retrieved from https://www.neps-data.de/Portals/0/Survey%20Papers/SP_XLIV.pdf