NEPS
National Educational Panel Study

Research Data

Starting Cohorts 3 and 4: 5th/9th Grade (SC3/SC4)
SUF Version 1.0.0
Data Manual [Supplement]: Weighting

Christian Aßmann, Hans Walter Steinhauer, and Sabine Zinn

SPONSORED BY THE

Federal Ministry of Education and Research

University of Bamberg, National Educational Panel Study (NEPS), 96045 Bamberg
https://www.neps-data.de
Principal Investigator: Prof. Dr. Hans-Günther Roßbach
Vice Managing Director: Prof. Dr. Sabine Weinert
Executive Director of Research: Dr. Jutta von Maurice
Executive Director of Administration: Dipl. sc. pol. Univ. Dipl.-Betriebswirt (FH) Gerd Bolz
Bamberg, 2012

Weighting the Fifth and Ninth Grader Cohort Samples of the National Educational Panel Study, Panel Cohorts

Technical Report

Christian Aßmann, Hans Walter Steinhauer and Sabine Zinn*

National Educational Panel Study
University of Bamberg

version: October 9, 2012

*methods.neps@uni-bamberg.de

1 Structure of sample

Target persons of starting cohorts (SC) three and four are all children attending secondary school in fifth or ninth grade in Germany in school year 2010/2011. Access to this population is gained via the corresponding institutions. The set of secondary schools involves all officially recognized and state-approved educational institutions in Germany providing schooling for fifth and/or ninth grade students. Children attending vocational schools or schools with a predominant foreign teaching language hindering the realization of a complete survey procedure with the available test instruments are excluded. Also, children not able to follow the normal testing procedure attending regular schools are excluded. Additionally the National Educational Panel Study (NEPS) comprises a sample of children attending special need schools. Here the focus is on schools with main emphasis on learning disabilities. Access to this population is gained via special need schools with federal-state-specific provisions explicitly for children with learning disabilities. Overall, 80% of children attending special need schools have a diagnosed learning disability – constituting the largest group of children in these schools. For more details also see Aßmann et al. (2011).

The samples for students in fifth and ninth grade are established on the basis of stratified multistage sampling designs utilizing a partly overlapping school sample. As common for surveys focusing on these age cohorts (e.g. PISA, TIMSS, TREE, etc.) the sample is drawn using an available sampling frame. On first stage schools are sampled and on second stage two classes (if available) of grade five and nine, respectively, are selected. Section 2 first describes the sampling of schools and classes, followed by the derivation of design weights. In section 3 a sample specific response propensity analysis is given. Section 4 gives insights on the trimming procedure applied to adjusted weights. Finally section 5 ends with recommendations on the use of weights.

2 Sampling and design weights

The large variety of federal-state-specific school systems is a challenge for sampling fifth and ninth grade children. Many different school types related to different transitions between elementary and secondary schooling institutions make up the set of schools providing access to the target population of fifth and ninth grade students. To reflect this large variety, seven explicit strata have been defined to sample schools. The first stratum comprises all Gymnasien (stratum GY), the second stratum consists of all Hauptschulen (stratum HS), the third stratum refers to all Realschulen (stratum RS), the fourth to Integrierte Gesamtschulen (stratum IG), the fifth includes schools offering all tracks of secondary education except the academic track (Schulen mit mehreren Bildungsgängen, stratum MB). The sixth explicit stratum comprises schools offering schooling to special needs children with learning disabilities (stratum FS). The seventh explicit stratum comprises all schools providing schooling only to fifth grade students, but not to ninth grade students (stratum $N5$). The definition of these seven explicit strata allows serving two important aspects of NEPS. First, a requisite of NEPS is to establish a sample of ninth grade students as the starting point of a longitudinal survey of young adults entering vocational education within the next years. In order to ensure sufficient sample sizes for statistical analysis within this heterogeneous population stemming to a large extent from Hauptschulen, Gesamtschulen and Schulen mit mehreren Bildungsgängen, NEPS comprises an oversampling of grade nine students attending these school types. The second aspect is reaching fifth and ninth grade students via the same set of schools, and thus reducing administrative survey costs. In addition to the explicit stratification according to school types, an implicit stratification based on federal states, regional classification and organizing institution has been used. Given the first stage sample of regular schools, at the second stage two school classes within each school are sampled randomly, if at least three classes are present, otherwise all classes are surveyed. In special need schools there is a census for all students. Subsequently, the derivation of design weights of students of grade nine and grade five is given.

Derivation of design weights of ninth grade students The frame for sampling schools providing access to the ninth grade children in school year 2010/11 has been set in school year 2008/09. Based on this frame, the sample of regular schools corresponding to the strata GY, HS, IG, MB and RS is established via systematic proportional to size (pps) sampling. The measure of size (mos) has been chosen proportionally to the number of classes of seventh grade in 2008/09 as the best available proxy for the number of classes in the ninth grade two years later (Aßmann et al., 2012). Based on extensive simulation studies, 629 out of a total of 11570 regular schools with seventh classes have been found sufficient to provide intended sample sizes addressing the above defined aspects (Aßmann et al., 2012). Special need schools corresponding to the stratum FS have been selected pps with measure of size proportional to the squared number of children attending grade nine. This measure of size allows handling the trade off between the expected sample size in this stratum and the total number of special need schools to be sampled. Their population consists of 1488 schools. Subsequently, M_h^9 denotes the total number of schools in stratum h and m_h^9 denotes the number of sampled schools in stratum h, $h \in \{GY, HS, IG, MB, RS, FS\}$. The measure of size for sampling a school j in stratum h is defined as

$$mos_{j\in h}^9 = \begin{cases} \frac{C_{j\in h}^7}{\min\{C_{j\in h}^7;2\}}, & \text{if } h \in \{GY, HS, IG, MB, RS\}, \\ \\ (S_j^9)^2, & \text{if } h \in \{FS\}. \end{cases}$$

Thereby $C_{j\in h}^7$ denotes the number of seventh classes which school j in stratum h hosts in school year 2008/09 and S_j^9 denotes the number of pupils attending a ninth grade in special need schools.[1] The strata specific total measure of sizes are

$$MOS_h^9 = \sum_{j\in h}^{M_h^9} mos_{j\in h}^9, \text{ for } h \in \{GY, HS, IG, MB, RS, FS\}.$$

[1] In most special need schools children in grade seven, eight and nine are educated together. That is, in the majority of cases no number of children in grade nine can be reported, but the total number of students in the grades seven to nine. Therefore, the number of ninth grade students is approximated by one third of the reported number of students in the grades seven to nine.

The values of M_h^9, m_h^9, mos_h^9, and MOS_h^9 are given in Table 1 for each considered stratum. While in special need schools all children are asked to participate, a subsample of classes in drawn in regular schools, i.e. if in a school j in stratum h the number of classes in the ninth grade $C^9_{j\in h}$ in school year 2010/11 exceeds two, a random sample of two ninth classes $C^9_{j\in h}$ has been drawn, otherwise all available classes have been selected. This procedure yields the following inclusion probabilities $\pi^9_{i\in j\in h}$ for a student i in secondary school j in stratum h, i.e.

$$\pi^9_{i\in j\in h} = \begin{cases} m_h^9 \cdot \frac{mos^9_{j\in h}}{MOS^9_h} \cdot \frac{\min\{C^9_{j\in h};2\}}{C^9_{j\in h}}, & \text{if } h \in \{GY, HS, IG, MB, RS\}, \\ m_h^9 \cdot \frac{(S_j^9)^2}{MOS^9_h}, & \text{if } h \in \{FS\}. \end{cases} \tag{1}$$

Derivation of design weights of fifth grade students The sample of schools providing access to the cohort of fifth graders accounts for the federal state specific timing of transition between primary and secondary schooling via explicit stratification. The sample of fifth grade students is based on all seven strata GY, HS, IG, MB, RS, FS and $N5$. Note that schools in stratum $N5$ are mainly primary schools educating students in classes one to six in Berlin and Brandenburg as well as schools in Hesse and Hamburg educating students in grades five and six only (schulformunabhängige Orientierungsstufen). To ensure ex ante least variation in design weights across the considered school strata, the number of schools that should be sampled per stratum has been calibrated via simulation studies. In total 240 regular schools and 65 special need schools have been found sufficient to reach the sample sizes of five graders intended by NEPS. More details on the simulation studies performed are given in Aßmann et al. (2012). To reduce administrative survey costs, for the strata GY, HS, RS, IG, MB and FS the fifth grade school sample of NEPS has been established as a subsample of the already realized ninth grade school sample. In stratum FS, subsampling has been performed via simple random sampling (srs). A srs design has been found sufficient because in the already sampled schools the high number of nine graders suggested also a high number of fifth graders.

Schools in stratum $N5$ contained in the frame referring to school year 2008/09 are selected randomly via systematic pps sampling. In the following, M_h^5 denotes the stratum specific total number of schools with fifth classes considered for sampling and m_h^5 denotes the stratum specific number of schools in the fifth grade sample. For the stratum $N5$, the measure of size $mos^5_{j\in N5}$ for sampling a school j is computed similarly to the measure of size of a comparable regular school in the ninth grade sample. It is given as

$$mos^5_{j\in N5} = \frac{C^5_{j\in N5}}{\min\{C^5_{j\in N5}; 2\}},$$

where $C^5_{j\in N5}$ is the number of fifth classes which school j in stratum $N5$ hosts in school year 2008/09. For regular a school j in stratum h, $h \in \{GY, HS, IG, MB, RS\}$, the measure of size is

$$mos^5_{j\in h} = \frac{\frac{C^5_{j\in h}}{\min\{C^5_{j\in h};2\}}}{\frac{C^7_{j\in h}}{\min\{C^7_{j\in h};2\}}},$$

where $C^5_{j\in h}$ is the number of fifth classes and $C^7_{j\in h}$ the number of seventh classes in school j in stratum h in school year 2008/09. The corresponding stratum specific total measure of size is

$$MOS^5_h = \sum_{j\in h}^{M^5_h} mos^5_{j\in h}, \quad \text{for} \quad h \in \{GY, HS, IG, MB, RS, N5\}.$$

For the fifth grade sample, the strata specific values of M^5_h, m^5_h,mos^5_h, and MOS^5_h are given in Table 1. For the strata $h \in \{GY, HS, IG, MB, RS, N5\}$, sampling of fifth grade students was performed similarly to the sampling of students in regular schools of the ninth grade sample: If in a sampled school at least three fifth classes were available in school year 2010/11, a random sample of two fifth classes has been drawn; otherwise the available number of classes has been selected. In conclusion, the inclusion probability $\pi^5_{i\in j\in h}$ of a fifth grader i in school j in stratum h is

$$\pi^5_{i\in j\in h} = \begin{cases} m^5_h \cdot \frac{mos^5_{j\in N5}}{MOS^5_{N5}} \cdot \frac{\min\{C^5_{j\in N5};2\}}{C^5_{j\in N5}}, & \text{if } h \in \{N5\} \\ \pi^9_{j\in h} \cdot \frac{m^5_h}{m^9_h}, & \text{if } h \in \{FS\} \\ \pi^9_{j\in h} \cdot m^5_h \cdot \frac{mos^5_{j\in h}}{MOS^5_h} \cdot \frac{\min\{C^5_{j\in h};2\}}{C^5_{j\in h}}, & \text{if } h \in \{GY, HS, IG, MB, RS\} \end{cases} \tag{2}$$

Here $\pi^9_{j\in h}$ denotes the probability that a school j in stratum h is part of the ninth grade sample, $h \in \{GY, HS, IG, MB, RS, FS\}$. The corresponding design weight of a fifth grade student i in school j in stratum h is then: $1/\pi^5_{i\in j\in h}$.

Note that the sample of fifth graders is enriched by a supplementary 214 cases related to an oversampling of fifth grade students having a Turkish migration background or a migration background related to the former Soviet Union. As no frame information is available allowing for direct identification, a two stage random selection process has been established. At the first stage, 500 schools from ten groups of federal states (50 schools per group of federal states) are sampled proportional to the number of students with foreign citizenship attending grade five. For this sample of 500 schools, the educational ministries of the federal states were asked to quantify the number of pupils related to the two migration backgrounds of interest. This quantification can take different forms based on the amount of information available in the federal state specific school statistics. For some federal states, the number of pupils with migration background is available directly, other federal states provided a top down listing of schools attended by the highest number of students with the two intended migration backgrounds. Based on this informational content a five categorical ordinal scale for the measure of size has been defined for a pps sampling of schools with a high number of Turkish migrants as well as those having a migration background from the former Soviet Union. In these schools all students with the Turkish migration background or a migration background related to the former Soviet Union have been asked to participate.

3 Weight adjustment for panel participation and non- response

Since the participation is not mandatory for schools or students they can both refuse to participate. This two stage decision process is reflected in the weighting strategy by adjusting the design weights using available information. With respect to school sampling information on school type, federal state, funding, regional classification, number of classes, expected number of students in the cohort, and school size are available, as well as some information arising during the process of school recruitment. The weights for the children are adjusted using information on the gross sample that were provided by school teachers.

When analyzing school participation, the replacement rule implemented to address school non-response in advance has to be considered, see Aßmann et al. (2011). Most often the schools refused participation in order to avoid further workload arising from participation in other studies. In order to counteract against the resulting sample size reduction on the level of students as much as possible, replacement schools have been defined in advance according to a matching rule. This matching rule defines replacement schools as schools from the same explicit stratum and similar implicit strata as the original school. However, the implemented replacement strategy was not able to completely hinder non-participation of schools.

Whether the available variables influence school participation is checked via probit regressions. To model the school participation of all contacted schools, their participation status is regressed on explaining factors reflecting the response burden for schools, the efforts involved in recruiting schools and all variables defining the considered explicit stratum and implicit strata. The effort of recruiting schools is measured by the number of schools contacted per state. A corresponding dummy variable separates the efforts by the median. As the legal basis for school participation differs across federal states, a federal state specific random effect is considered. Hence, the regression is given as

$$y_{j\in h} = \begin{cases} 1, & \text{if } y^*_{j\in h} > 0 \\ 0, & \text{if } y^*_{j\in h} \leq 0 \end{cases} \quad \text{with} \quad y^*_{j\in h} = \alpha + X_j\beta + Z_h\gamma + \xi_h + \epsilon_{j\in h},$$

where $j \in h$ indicates school j located in stratum h, and $\xi_h \sim \mathcal{N}(0, \sigma^2_\xi)$ and $\epsilon_{j\in h} \sim \mathcal{N}(0, 1)$. For schools not related to the migrants oversampling, the estimated random intercept models are shown in Table 2. The results indicate the significance of variables defining the explicit and implicit strata.[2] The results of the regression model describing the participation propensity of schools contacted during the migrants oversampling are given in Table 3.

For the model estimating the participation of special need schools the dummy for the effort measurement is significant. For the schools sampled in the oversampling of students with migrational background this dummy could not be formed. Furthermore, here the state specific random effect is not significant. For each of the strata models for the regular

[2] Note, that some variables are stratum specific and can only be considered in certain models, e.g. number of classes in grade seven cannot be considered in stratum $N5$. Besides, collinearity within the $N5$ stratum does not allow for the consideration of the variable effort made in recruitment in the associated stratum specific model.

schools as well as for special need schools the effort made in school recruitment (if in the model) is significant. In case of Hauptschulen ($h = HS$) the dummy indicating public funding is significant. Therefore, the weights in the corresponding samples of schools need to be adjusted by states membership and in case of Hauptschulen they are additionally adjusted for "funding". The weight of school j in state f and stratum h is adjusted by the factor

$$\gamma_{jfh} = \frac{m^r_{fh} + m^n_{fh}}{m^r_{fh}},$$

where m^r_{fh} denotes the number of participating schools and m^n_{fh} the number of non-participating schools in state f in stratum h. The adjustment on the school level is only done for the first wave since the participation of schools is the necessary condition to get access to the students on the second stage. Having access to the sample population the decision to participate in the survey is on the students level. However, since most students mostly are not of legal age, they need the permission of their parents. In the first wave the decision of a student to participate in the survey is conditional on the schools participation. Contrasted by future waves where a student can participate even if the school refuses to participate.

After adjusting weights on the first stage the students weights on the second stage need to be adjusted as well. The response propensity of students is modelled via probit models regressing the participation status on characteristics available for the gross sample of students. Starting cohort specific models are estimated, with a further distinction between students attending regular and special need schools. Thereby the different sets of information available for the different starting cohorts is taken into account. For example, for the ninth graders last year's school grades are provided by teachers, which are not necessarily available for fifth graders. For starting cohort three a separate model for the groups of oversampled migrants is estimated. In addition to the information available on student characteristics, random intercepts are considered to reflect the cluster structure of the gross samples of children. The complete models by starting cohorts and samples are given in Table 4.

For starting cohort three the models are estimated for the sample of students in regular and special need schools as well as for the sample of students with migrational background. For all three samples the propensity to participate is significantly negatively influenced by another language than German spoken at home as well as by missing values in characteristics related to competencies (i.e. competencies in maths and reading, dyslexia, special educational needs and attention deficit hyperactivity disorder). For students in regular schools the participation propensity is positively influenced by good competencies in maths and reading and negatively influenced by missing values in personal attributes (i.e. gender, year of birth) and migration characteristics (i.e. language spoken at home, nationality). Since competencies in maths and reading are subject to analysis, for further studies this relationship should be minded to avoid potential bias.

For all considered five subgroups of the starting cohorts (fifth and ninth graders in regular schools and special need schools, as well as the migrants oversampling of fifth graders) some of the findings can be generalized. In each subgroup the participation propensity of the students is effected in the same way: male students and students speaking another language than German at home have a lower propensity to participate. Further missing values in personal, migrational and competence characteristics are a strong predictor influencing the participation negatively. A positive effect is found for good grades in maths and, except for the students with migrational background, younger students are more willing to participate.

Students are educated in different school systems that vary across states. To categorize the diversity of parallel existing school systems within the states we use the approach developed by Tillmann (2012). The categorization yields a negative effect for binomial and quadrinomial school systems whereas only the latter is significant. For students in special need schools the maths competence has a significant positive effect on the participation propensity and for students with a migrational background the dummy indicating a Turkish migrational background is highly significant and influences the response propensity negatively. The size of the test groups yield different effects throughout the three samples. While being insignificant in regular schools it is significantly negative for special need schools and positive for students with a migrational background.

In starting cohort four the effects of gender, grade in maths and the missing values in competence characteristics are of similar kind. Male students and students with at least half of the competence characteristics missing are less willing to participate whereas students with a good grade in maths are more willing. For students in regular schools the younger students are significantly more willing to participate, also good grades in German have a positive effect on the participation propensity. Surprisingly, non-German students have a higher propensity to participate in the grade nine sample than German students. Negative effects are found for students with special educational needs and for those with missing values in personal characteristics.

4 Weight trimming

Aiming at an increased statistical efficiency of weighted analysis the adjusted design weights have been trimmed. The general goal of weight trimming is to reduce the sampling variance and at the same time to compensate for potential increase in bias. Trimming has been performed using the so called “Weight Distribution” approach (Potter, 1990). Here, design weights are assumed to follow an inverse beta distribution with a cumulative distribution function F_w. Parameters of the sampling weight distribution are estimated using the sampling weights and a trimming level τ is computed whose occurrence probability is one percent, i.e., $1 - F_w(\tau) = 0.01$. Sampling weights in excess of τ are trimmed to this level and the excess is distributed among the untrimmed weights. The parameters for the sampling weight distribution are then again estimated using the trimmed adjusted weights and a revised trimming level $\tilde{\tau}$ is computed. The trimmed adjusted weights are compared to the revised level $\tilde{\tau}$. If any weights are in excess of $\tilde{\tau}$, they are trimmed to this level and the excess is distributed among the untrimmed weights. This procedure is iteratively repeated until no weights are in excess of a newly revised trimming level. To ease statistical analysis the trimmed design weights are standardized with mean one. Subsequently, their distribution is summarized (abbr.: REG - regular schools, FS - special need schools, MIG - migrants oversampling):

Starting cohort	Number of students	Min.	Lower Quart.	Median	Mean	Upper Quart.	Max.
SC 4	16425	0.05	0.76	1.01	1.00	1.20	2.40
REG	15239	0.16	0.80	1.03	1.04	1.21	2.40
FS	1186	0.05	0.15	0.23	0.43	0.42	2.40
SC 3	6112	0.00	0.46	0.69	1.00	1.15	4.49
REG	5283	0.14	0.51	0.72	1.00	1.14	4.49
FS	587	0.02	0.04	0.09	0.72	0.31	4.49
MIG	242	0.00	0.03	0.58	1.61	3.86	4.49

5 Use of weights

Given the quite complex structure of the samples of the secondary school students, no final recommendation are at hand concerning the use of design and adjusted weights. In general, the use of design weights is recommended when analysis is concerned with sample data stemming from different strata – presumably the case in most analysis. Although, there are no general results available how the use of design or calibrated weights render any possible analysis (see Rohwer (2011) for a general discussion) the use of weights may possibly help to highlight important features of the analysis under consideration not at least serving as a robustness check for the performed analysis. Adjusted design weights provided are labeled as `weight_design` and standardized design weights are labeled as `weight_design_std`. The subsequent table lists all types of weights provided:

Type of weight	**Label**
adjusted design weight	`weight_design`
trimmed design weight, standardized with mean one	`weight_design_std`

References

Aßmann, C., Steinhauer, H. W., Kiesl, H., Koch, S., Schönberger, B., Müller-Kuller, A., … Blossfeld, H.-P. (2011). Sampling designs of the National Educational Panel Study: challenges and solutions. In H.-P. Blossfeld, H.-G. Roßbach, & J. v. Maurice (Eds.), *Education as a lifelong process* (Vol. 14, pp. 51–65). Wiesbaden: VS Verlag für Sozialwissenschaften. doi: 10.1007/s11618-011-0181-8

Aßmann, C., Steinhauer, H. W., & Rässler, S. (2012). Aspekte der Stichprobenziehung in der erziehungswissenschaftlichen Forschung. In S. Maschke & L. Stecher (Eds.), *Enzyklopädie der Erziehungswissenschaft Online (EEO), Fachgebiet Methoden der empirischen erziehungswissenschaftlichen Forschung* (pp. 1–15). Weinheim und Basel: Beltz Juventa. doi: 10.3262/EEO07120215

Potter, F. J. (1990). A study of procedures to identify and trim extreme sampling weights. In American Statistical Association (Ed.), *Proceedings of the Survey Research Methods Section* (pp. 225–230). Retrieved 25.09.2012, from `http://www.amstat.org/sections/srms/Proceedings/papers/1990_034.pdf`

Tillmann, K.-J. (2012). Das Sekundarschulsystem auf dem Weg in die Zweigliedrigkeit: Historische Linien und aktuelle Verwirrungen. *Pädagogik*, *64*(5), 8–12.

Table 1: Population sizes, sample sizes and measures of size by strata.

	SC4			SC3		
Stratum	m_h^9	M_h^9	MOS_h^9	m_h^5	M_h^5	MOS_h^5
N5	0	0	0	26	1383	1914.50
GY	154	2970	5589.5	214	480*	559.06
HS	233	3990	4641.5			
IG	70	822	1727.5			
MB	64	1288	1524.0			
RS	108	2500	3936.5			
FS	110	1488	489128.264	65	110	——

* For $h \in \{GY, HS, IG, MB, RS\}$ the population of 480 schools with fifth classes is a sub-sample of the realized grade nine school sample.

Table 2: Results of random intercept models for school participation (by strata). Standard deviations are given in parentheses.

	Stata						
	N5	GY	HS	IG	MB	RS	FS
constant	−1.283	−1.065	−1.868***	−0.275	−0.799	−0.999	0.137
	(1.041)	(0.598)	(0.473)	(0.529)	(1.095)	(0.714)	(0.224)
number of students	0.001	0.001	0.001	0.001	−0.002	−0.000	
in school	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	
number of students	0.010	−0.001	−0.001	−0.002	0.000	−0.004	
in age group	(0.020)	(0.002)	(0.002)	(0.003)	(0.005)	(0.003)	
number of students							−0.002
in grade nine (squared)							(0.021)
organizing institution	0.762	0.127	1.043**	−0.063	−0.348	−0.012	
public	(0.662)	(0.207)	(0.372)	(0.377)	(0.668)	(0.304)	
urbanization	−0.027	0.233	0.164	−0.179	−0.132	0.088	−0.341
rural	(0.661)	(0.287)	(0.168)	(0.459)	(0.365)	(0.295)	(0.311)
urbanization	−0.790	−0.069	0.085	−0.242	−0.430	−0.215	−0.138
urban	(0.543)	(0.139)	(0.123)	(0.294)	(0.342)	(0.158)	(0.201)
number of classes	−0.020	0.025	−0.038	−0.026	−0.107	−0.014	
in grade five	(0.087)	(0.041)	(0.057)	(0.051)	(0.130)	(0.054)	
number of classes		0.021	0.025	−0.079	0.137	0.108	
in grade seven		(0.069)	(0.032)	(0.048)	(0.078)	(0.096)	
cohorts sampled							−0.082
five and nine							(0.183)
effort in recruitment							0.708**
up to 8							(0.253)
effort in recruitment		0.807***	0.435	1.011**	1.192**	0.295	
up to 51		(0.220)	(0.478)	(0.317)	(0.418)	(0.538)	
random intercept	0.000	0.007	0.065	0.042	0.144	0.021	0.000
federal states	(0.000)	(0.084)	(0.256)	(0.205)	(0.379)	(0.145)	(0.000)
schools per stratum	54	415	672	218	110	337	203

Notes: ***, ** and * denote significance at the 0.1%, 1% and 5% level respectively.

Table 3: Results of random intercept model for the participation of schools contacted during the migrants oversampling. Standard deviations are given in parentheses.

Variable	Estimated coefficient
constant	-1.538
	(0.798)
number of students	0.001
in school	(0.001)
school type	-0.039
GY	(0.695)
school type	0.407
HS	(0.657)
school type	-0.368
IG	(0.711)
school type	0.951
MB	(0.737)
school type	-0.327
RS	(0.669)
stratum	-0.400
Russian	(0.383)
stratum	0.193
Turkish	(0.322)
urbanization	-4.502
rural	(146.097)
urbanization	-0.008
urban	(0.437)
random intercept	0.223
federal states	(0.472)
number of schools	198

Table 4: Results of response propensity models for starting cohorts three and four (REG - regular schools, FS - special need schools, MIG - oversampling of migrants). Standard deviations are given in parentheses.

	SC3 - REG	SC3 - FS	SC3 - MIG	SC4 - REG	SC4 - FS
constant	0.859^{*}	1.521^{**}	0.308	0.546	1.166^{**}
	(0.410)	(0.483)	(0.257)	(0.437)	(0.365)
testgroup size	−0.021	-0.052^{*}	0.013^{**}	−0.005	−0.025
	(0.014)	(0.023)	(0.004)	(0.007)	(0.017)
gender	−0.005	-0.356^{*}	−0.162	-0.145^{***}	-0.304^{**}
male	(0.048)	(0.159)	(0.104)	(0.029)	(0.106)
language spoken at home	-0.288^{**}	-0.988^{***}	-0.451^{***}	−0.104	−0.211
no German	(0.101)	(0.289)	(0.134)	(0.057)	(0.198)
nationality	−0.116	0.451	0.148	0.123^{*}	0.060
not German	(0.118)	(0.306)	(0.128)	(0.061)	(0.199)
dyslexia	0.040	0.444	0.607	0.027	0.302
yes	(0.120)	(0.402)	(0.535)	(0.082)	(0.260)
grade in maths	0.310^{***}	0.535^{**}	0.223	0.144^{***}	0.472^{***}
1 to 3	(0.066)	(0.197)	(0.128)	(0.031)	(0.135)
reading competence	0.243^{***}	−0.006	0.052		
1 to 3	(0.070)	(0.191)	(0.130)		
grade in German				0.144^{***}	−0.064
1 to 3				(0.034)	(0.140)
year of birth	0.093	0.219	−0.101		
after 1999	(0.050)	(0.184)	(0.117)		
year of birth	−0.130	0.227	0.278		
before 1999	(0.111)	(0.208)	(0.192)		
year of birth				0.112^{***}	0.267
after 1995				(0.032)	(0.141)
year of birth				−0.086	0.018
before 1995				(0.044)	(0.123)
special educational needs	0.102		−0.195	-0.586^{***}	
yes	(0.188)		(0.270)	(0.133)	
migrational background			-0.909^{***}		
Turkish			(0.179)		
school type	−0.014				
GS	(0.398)				
school type	0.264			0.470	
GY	(0.281)			(0.474)	
school type	−0.095			0.309	
HS	(0.305)			(0.477)	
school type	−0.142			0.232	
IG	(0.360)			(0.495)	
school type	0.212			0.495	
MB	(0.402)			(0.505)	
school type	0.167			0.216	
RS	(0.284)			(0.476)	
school system	−0.396			−0.036	
binomial extended	(0.256)			(0.263)	
school system	−0.677			−0.097	
binomial pure	(0.389)			(0.326)	
school system	-0.424^{*}			−0.070	
quadinomial extended	(0.187)			(0.257)	
missings in personal characteristics	-4.110^{***}			-2.651^{***}	
yes	(0.549)			(0.193)	
missings in migration characteristics	-1.255^{***}	−0.332	0.787^{**}	−0.075	−0.356
yes	(0.228)	(0.703)	(0.304)	(0.092)	(0.425)
missings in competence characteristics	-1.522^{***}	-4.371^{***}	-2.849^{***}	-0.773^{***}	-2.880^{***}
at least half ot them	(0.251)	(1.258)	(0.361)	(0.174)	(0.429)
attention deficit hyperactivity disorder		−0.117			0.354
yes		(0.246)			(0.226)
random intercept	0.969	1.236	0.275		1.540
school level	(0.984)	(1.112)	(0.524)		(1.241)
random intercept				0.156	
school type and state				(0.395)	
random intercept				0.073	
testgroup size				(0.395)	
participants + non-participants	9081	962	836	23327	2146

Notes: ***, ** and * denote significance at the 0.1%, 1% and 5% level respectively.