NEPS
National Educational Panel Study

NEPS Working Papers

Sabine Zinn

Replication Weights for the Cohort Samples of Students in Grade 5 and 9 in the National Educational Panel Study

NEPS Working Paper No. 27

Bamberg, July 2013

SPONSORED BY THE
Federal Ministry of Education and Research

Working Papers of the German National Educational Panel Study (NEPS)
at the University of Bamberg

The NEPS Working Papers publish articles, expertises, and findings related to the German National Educational Panel Study (NEPS).
The NEPS Working Papers are edited by a board of researchers representing the wide range of disciplines covered by NEPS. The series started in 2011.

Papers appear in this series as work in progress and may also appear elsewhere. They often represent preliminary studies and are circulated to encourage discussion. Citation of such a paper should account for its provisional character.

Any opinions expressed in this series are those of the author(s) and not those of the NEPS Consortium.

The NEPS Working Papers are available at
http://www.uni-bamberg.de/neps/publikationen/neps-working-papers/

Contact: German National Educational Panel Study (NEPS) − University of Bamberg − 96045 Bamberg − Germany − contact.neps@uni-bamberg.de

Replication Weights for the Cohort Samples of Students in Grade 5 and 9 in the National Educational Panel Study

Sabine Zinn, University of Bamberg

E-mail address of lead author:
sabine.zinn@uni-bamberg.de

Bibliographic data:

Zinn, S. (2013). *Replication weights for the cohort samples of students in Grade 5 and 9 in the National Educational Panel Study* (NEPS Working Paper No. 27). Bamberg: University of Bamberg, National Educational Panel Study.

Replication Weights for the Cohort Samples of Students in Grade 5 and 9 in the National Educational Panel Study

Abstract

In order to obtain valid inference, the analysis of survey data requires special approaches to account for sampling design features. This is particularly true for analyzing complex survey data in which inclusion probabilities are not constant—as is the case, for example, in the National Educational Panel Study (NEPS). One way to achieve proper results—even when a statistical method does not explicitly account for survey design features—is the method of balanced repeated replication. This methodology provides a correct assessment of the variances for a wide range of estimators from stratified multistage sampling designs. Balanced repeated replication can be applied without further ado if so-called replication weights are available. To facilitate an unbiased variance estimation for NEPS data users, the NEPS methods group provides specific replication weights for the students participating in the first wave of the NEPS fifth-grade sample and for students participating in the first and the second wave of the NEPS ninth-grade sample as well. Additionally, replication weights are provided for fifth and ninth graders, for whom an interview with one parent could be realized. In this paper we describe how these weights have been derived and how they can be used to yield valid variance estimates.

Keywords

variance estimation, method of balanced repeated replication, National Educational Panel Study, fifth- and ninth-grade sample

1 Introduction

The samples of the fifth-grade cohort (Starting Cohort 3, SC3) and the ninth-grade cohort (Starting Cohort 4, SC4) of the National Educational Panel Study (NEPS) have been established using stratified multistage sampling. Statistical standard estimation techniques might face difficulties in handling such a design because they are often predicated on simple random sampling. Regardlessly applying them to a multistage sampling design increases the risk of underestimating the variability of survey statistics. Here, the method of balanced repeated replication (BRR) offers a remedy. It provides a correct assessment of the variances for a wide range of estimators from stratified multistage sampling designs. Examples of such estimators are ratios of two estimated totals, correlation coefficients, and regression coefficients (Shao, 1996). The basic idea of the BRR is to construct a set of replicates consisting of random groups that represent the sampling units in the different strata. Forming these groups in a balanced way allows us to compute unbiased variance estimates (Wolter, 2007; Rust & Rao, 1996). Balanced repeated replication can be applied without further ado if so-called replication weights are available. Replication weights make it possible to derive the set of replicates necessary for variance estimation by simply multiplying them with the parent sample. To facilitate an unbiased variance estimation for NEPS data users, the NEPS methods group provides specific replication weights for the students participating in the first wave of SC3 as well as for students participating in the first and the second wave of SC4. Additionally, replication weights are provided for fifth and ninth graders for whom an interview with one parent could be realized. The replication weights are part of the Scientific Use Files (SUF) for SC3 and SC4 Version 1.1.0.[1] This paper describes how these replication weights for the BRR have been derived.[2] The remainder of the paper is organized as follows: In Section 2, we briefly describe the structure and the sampling design of SC3 and SC4. Section 3 presents the concept of the method of balanced repeated replication. Section 4 illustrates the derivation of the corresponding replication weights for SC3 and SC4. In conjunction with this, we give details on the adjustment of weights necessary to concord with the design of the parent sample. Finally, Section 5 gives a summary of the provided sets of replication weights.

2 Sampling Design of the Cohort Samples of Students in Grade 5 and 9

The fifth- and ninth-grade sample of NEPS comprises children attending secondary school in the Grade 5 and 9 in Germany in the school year of 2010/2011. The sample has been built upon a stratified multistage sampling design (Aßmann et al., 2011, 2012). Schools were selected from the set of officially recognized and state-approved secondary schools in Germany. In this process, six different school types were differentiated: Special schools, *Gymnasien*, *Hauptschulen*, *Realschulen*, *Integrierte Gesamtschulen*, and schools offering all tracks of secondary education except the academic track (in the following denoted as 'MB'). Subsequently, we refer to the latter five school types as regular schools. Special schools and regular schools form the six explicit strata of the ninth-grade sample. The fifth-grade sample consists of three explicit strata, partly overlapping with the ninth-grade sample. That is, the first explicit stratum of the fifth-grade sample has been established on the basis of five of the six explicit strata of the ninth-grade sample. In more detail, the stratum comprises fifth graders from regular schools that provide schooling to students in Grade 5 and 9. Special schools make up the second explicit stratum of the fifth-grade sample. Finally, the third explicit stratum of the fifth-grade sample contains children attending schools that provide schooling to fifth graders only and not to ninth graders.

[1] doi 10.5157/NEPS:SC3:1.1.0 and doi 10.5157/NEPS:SC4:1.1.0

[2] The paper is a condensed version of Zinn (2013, forthcoming).

Table 1: Stratum-specific numbers of sampled schools and students in Wave 1 and 2 of the Grade 9 sample.

Stratum	Schools	Students in Wave 1	Students in Wave 1 & 2
Gymnasien	149	5,118	5,069
Hauptschulen	181	3,570	3,515
Realschulen	104	3,108	3,069
Gesamtschulen	55	1,617	1,609
MB	56	1,127	1,116
Special schools	103	1,089	930

Table 2: Stratum-specific numbers of sampled schools and students in Wave 1 of the Grade 5 sample.

Stratum	Schools	Students
Regular schools offering schooling to fifth graders and not to ninth graders	21	430
Regular schools offering schooling to fifth graders and to ninth graders	182	4,559
Special schools	57	570
Schools holding students with a Turkish migration background or a migration background related to the former Soviet Union	31	219

After sampling schools on the first stage, two classes (if available) of Grade 5 and 9 were sampled respectively within regular schools on a second stage. After that, all children in the selected classes were asked for participation. In special schools simple random sampling was performed, that is, students of all classes were asked to participate in NEPS. An additional sample of children with a Turkish migration background or a migration background related to the former Soviet Union complements the fifth-grade cohort sample. The sample has been established by sampling relevant students from schools with a high number of predicted migrants (Aßmann et al., 2012). For the purposes of balanced repeated replication students that are part of this sample were taken to form a stratum of their own.

Overall, the ninth-grade sample comprises information of students from 648 schools: $15,629$ students participated in the first wave, and $15,308$ students participated in the first and second wave. The first wave of the fifth-grade sample comprises 291 schools and contains information on $5,778$ students. Table 1 and 2 show the number of schools and students according to the grade sampled within the different strata. The parents of all students who are participating in NEPS have also been asked for an interview. In total $8,677$ parents of ninth graders, and $3,974$ parents of fifth graders could be interviewed.

3 The Method of Balanced Repeated Replication

The method of balanced repeated replication is a widely-used technique for variance estimation in surveys that are subject to stratified multistage sampling. It was first introduced by McCarthy (1969) for a case where only two primary sampling units are sampled with replacement in the first sampling stage. Today several extensions to the original approach exist that allow the BRR to be applied to a wider scope of tasks, see, for example, Rao & Shao (1999); Shao & Chen (1999); Shao et al. (1998); Saigo et al. (2001). Before we explore the essentials of the BRR in more detail, we will first describe the statistical setting to which the BRR is applied.

3.1 The Setting

Let us assume that we face a survey sample subject to stratified multistage sampling involving H strata. Each stratum h comprises n_h primary sampling units (PSUs) and every primary sampling unit i contains $m_{(h,i)}$ secondary sampling units (SSUs) j. All units k being part of secondary sampling units are constituted to be fully sampled. If survey weights are available for all sampled elements, an unbiased estimator of a population total Y for a variable y is given by (Rao & Shao, 1999)

$$\hat{Y} = \sum_{(h,i,j,k)\in s} w_{hijk} y_{hijk},$$

where s describes the sample, y_{hijk} is the value of variable y associated to unit (h, i, j, k), and w_{hijk} the corresponding sampling weight. In many cases, a survey estimator $\hat{\theta}$ can be written as a function $g(\cdot)$ of a vector of estimated totals:

$$\hat{\theta} = g(\hat{A}),$$

with

$$\hat{A} = \sum_{(h,i,j,k)} w_{hijk} a_{hijk},$$

and a_{hijk} is a vector of values corresponding to unit (h, i, j, k). Examples of such estimators are ratios of two estimated totals, correlation coefficients, and regression coefficients (Shao, 1996). Assume, for instance, that we are interested in the prevalence of learning disabilities among male students. An estimator for this quantity is the ratio of the number of male students with learning disabilities to the number of male students. Thus, it can be expressed as

$$\hat{\theta} = g(\hat{A}) = \frac{\sum_{(h,i,j,k)} w_{hijk} z_{hijk}}{\sum_{(h,i,j,k)} w_{hijk} x_{hijk}},$$

where x_{hijk} is a dichotomous variable that is coded by 1 if a student is male and 0 otherwise, and z_{hijk} is a dichotomous variable that is 1 if a male student suffers from a learning disability and 0 otherwise.

3.2 The Method

The basic idea of the BRR is to construct a set of balanced replicates from random groups in the parent sample. Commonly, random groups are formed from the primary sampling units only, irrespective of any further subsampling. Such proceeding is predicated on the fact that the sampling variance can be approximated adequately from the variation between the totals of the primary sampling units when the first-stage sampling fraction is small (which is usually the case). In survey statistics this practice is known as the ultimate cluster approximation (Kalton, 1979; Lee & Forthofer, 2006). In its original version the BRR assumes only two primary sampling units per stratum, that is, $n_h = 2$ for all strata h. A single replicate is formed by systematically deleting one primary sampling unit from each stratum and then doubling the sampling weights of the remaining primary sampling units. Hence, the replication weight $w_{hijk}^{(r)}$ of entity k in SSU j located in PSU i in stratum h corresponding to the rth replicate is $((h, i, jk) \in s, r = 1, \ldots, R)$:

$$w_{hijk}^{(r)} = \begin{cases} 2w_{hijk}, & \text{if PSU } i \text{ from stratum } h \text{ is part of the } r\text{th replicate,} \\ 0, & \text{otherwise.} \end{cases}$$

Because of the practice to neglect half of the parent sample within each replicate, the BRR is also called the method of balanced half-samples. To promote unbiased variance estimators, the set of replicates has to be *balanced* (Wolter, 2007). That is, each pair of primary sampling units from different strata must have the same frequency of appearing in the set of replicates. This condition can be formalized to

$$\sum_{r=1}^{R} \delta_h^{(r)} \delta_k^{(r)} = 0 \quad \text{for all } h \neq k; h, k = 1, \ldots, H,$$

with

$$\delta_h^{(r)} = \begin{cases} +1, & \text{if PSU 1 from stratum } h \text{ is part of the } r\text{th replicate,} \\ -1, & \text{if PSU 2 from stratum } h \text{ is part of the } r\text{th replicate.} \end{cases}$$

A minimal set of balanced replicates can be derived using a Hadamard matrix[3] of order R where R is the smallest multiple of four, greater then H:

$$H \leq R \leq H + 3. \tag{1}$$

In more detail, the entries (h, r) of a Hadamard matrix A of order R determine the primary sampling units that have to remain in a half-sample to obtain a balanced set of replicates,[4] that is, $\delta_h^{(r)}$ equals entry (h, r) of matrix A. An unbiased variance estimate is then obtained by

$$var(\hat{\theta}) = \frac{1}{R} \sum_r (\hat{\theta}^{(r)} - \hat{\theta})(\hat{\theta}^{(r)} - \hat{\theta})',$$

where $\hat{\theta}^{(r)}$ is the survey estimate based on replicate r (i.e., weighted with the replication weights $w_{hijk}^{(r)}$).

One crucial requisite for the feasibility of any method of repeated replication is that each single set of replication weights has to maintain the representation of the population structure in the sample, i.e., replication weights have to be adjusted for unit nonresponse. At this point, the BRR might encounter a number of severe problems: As the method implies deleting half of the parent sample, very small sample sizes might result from this, causing unfeasible adjustment factors for nonresponse and, therefore, nonsensical replication weights. A simple variant of the BRR that allows us to overcome this difficulty is perturbing the replication weights by a factor ϵ, $\epsilon \in \{0, 1\}$ (Judkins, 1990):

$$w_{hijk}^{(r)}(\epsilon) = \begin{cases} (1+\epsilon) w_{hijk}, & \text{if PSU } i \text{ from stratum } h \text{ is in the } r\text{th replicate,} \\ (1-\epsilon) w_{hijk}, & \text{otherwise.} \end{cases} \tag{2}$$

The variance estimator resulting from this is

$$var(\hat{\theta}) = \frac{1}{\epsilon^2 R} \sum_r (\hat{\theta}^{(r)}(\epsilon) - \hat{\theta})(\hat{\theta}^{(r)}(\epsilon) - \hat{\theta})'. \tag{3}$$

For convenience, mostly ϵ is set to 0.5 (Rust & Rao, 1996).

The BRR can easily be extended to cases where strata comprise more than two primary sampling units. The basic idea is to randomly divide the set of primary sampling units in each stratum h into two groups of almost the same sizes g_h^1 and g_h^2, that is,

$$g_h^1 = [n_h/2] \text{ and } g_h^2 = n_h - g_h^1. \tag{4}$$

[3]A Hadamard matrix is a square matrix whose entries are either −1 or +1 and whose rows are mutually orthogonal.

[4]Here the row of a Hadamard matrix that contains only +1s has to be excluded.

By means of these groups a set of balanced replicates can still be constructed using Hadamard matrices (Rao & Shao, 1996): The entry (h, r) of a Hadamard matrix of order R determines whether Group 1 or Group 2 is assigned to a half-sample. Figure 1 illustrates this creation of replicates. For the computation of a survey estimate $\hat{\theta}^{(r)}$ based on replicate r, the replication

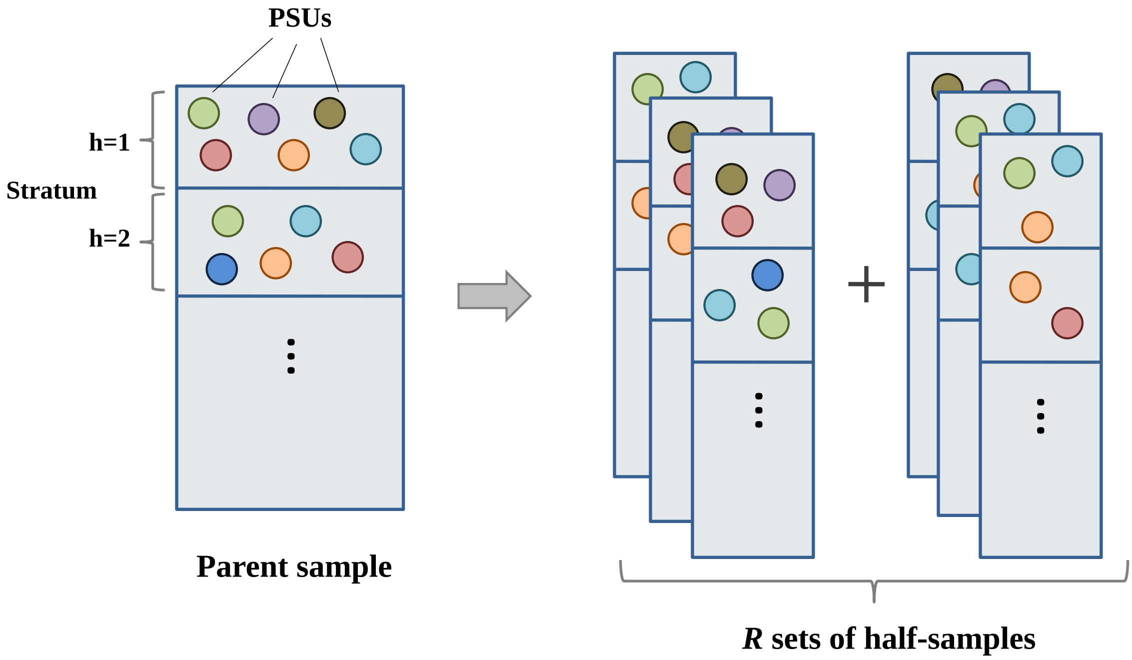

Figure 1: Creation of replicates by assigning for each stratum separately groups of primary units (PSUs) to half-samples.

weights (2) have to be modified in the following way

$$w_{hijk}^{(r)}(\epsilon) = \begin{cases} \left(1 + \epsilon\sqrt{\frac{n_h - g_h^1}{g_h^1}}\right) w_{hijk}, & \text{if entry (h,r) = +1,} \\ \left(1 - \epsilon\sqrt{\frac{g_h^1}{n_h - g_h^1}}\right) w_{hijk}, & \text{if entry (h,r) = -1.} \end{cases} \tag{5}$$

The variance estimator of this BRR variant does not change and is given by equation (3). In surveys with very large sample sizes this grouped variant of the BRR might produce asymptotically incorrect results (Valliant, 1987). To overcome this issue, Rao & Shao (1996) suggest repeating the random grouping T times and taking the average of the resulting T BRR variance estimators. To put it more succinctly, at first on the basis of always the same Hadamard matrix of order R from the primary sampling units in each stratum T times random groups are formed (resulting in R multiplied by T replicates). Then for the T sets of random groups variance estimators $var^t(\hat{\theta})$ are computed ($t = 1, \ldots, T$). The mean of these constitutes the revised variance estimator:

$$var_T(\hat{\theta}) = \frac{1}{T}\sum_{t=1}^{T} var^t(\hat{\theta}). \tag{6}$$

In simulation studies, Rao & Shao (1996) found $T = 30$ to produce unbiased results. The group variant of the BRR fits well to the design of the school samples of NEPS and facilitates computing unbiased variance estimators of related survey statistics.

The basic sampling weights might be subject to poststratification and unit nonresponse adjustment. To capture the possible impact of the weight adjustment on variance estimates, each set of replicates has to be treated with the same adjustment steps as applied to the sampling weights (Rao & Shao, 1999).

4 Construction of Replication Weights

The group variant of the BRR is well suited for estimating sample variances from the fifth- and ninth-grade cohort samples of NEPS. Its central piece is the stratum-wise formation of random groups of primary sampling units. In accordance with this, we have randomly divided the set of schools in each stratum into two groups of almost the same size. To determine the group sizes, we relied on formula (4). Once the groups had been created, we assigned them to one of the two half-samples of each replicate. For this purpose, we have used a Hadamard matrix A of order eight:

$$A = \begin{pmatrix} +1 & +1 & +1 & +1 & +1 & +1 & +1 & +1 \\ +1 & -1 & +1 & -1 & +1 & -1 & +1 & -1 \\ +1 & +1 & -1 & -1 & +1 & +1 & -1 & -1 \\ +1 & -1 & -1 & +1 & +1 & -1 & -1 & +1 \\ +1 & +1 & +1 & +1 & -1 & -1 & -1 & -1 \\ +1 & -1 & +1 & -1 & -1 & +1 & -1 & +1 \\ +1 & +1 & -1 & -1 & -1 & -1 & +1 & +1 \\ +1 & -1 & -1 & +1 & -1 & +1 & +1 & -1 \end{pmatrix}$$

This matrix allowed us to derive a minimal set of balanced replicates: For $H = 4$ (the number of strata in the fifth-grade sample) and for $H = 6$ (the number of strata in the ninth-grade sample) $R = 8$ satisfies the condition (1). The entries of A (except for the row containing only +1s) determine, on the one hand, to which half-sample a random group of schools will be assigned and, on the other, how many replicates will be created (here, eight). To give an example: The stratum *Gymnasien* of the ninth-grade sample comprises 149 schools, and the stratum *Hauptschulen* contains 181 schools, see Table 1. In each stratum, two random groups have been created: In the stratum *Gymnasien*, Group 1 includes 124 schools, and Group 2 contains 125 schools. In the stratum *Hauptschulen*, 140 schools have been assigned to Group 1 and 141 schools to Group 2. According to the corresponding entries of A (namely, $(2, 1) = +1$ and $(3, 1) = +1$), in the first replicate group, one of both strata is part of Half-Sample 1 and Group 2 is part of Half-Sample 2. By contrast, Half-Sample 1 of the second replicate includes Group 1 of the stratum *Gymnasien* and Group 2 of the stratum *Hauptschulen*. Accordingly, the second half-sample contains Group 2 of the stratum *Gymnasien* and Group 1 of the stratum *Hauptschulen*. The respective entries of A are $(2, 2) = +1$ and $(3, 2) = -1$. Figure 2 illustrates this example. Note that only so many rows of H are used as strata exist; all remaining rows remain untouched. Relying on the assignment of school groups to half-samples, formula (5) determines the computation of the corresponding replication weights. To improve the applicability of the BRR, we repeat the creation of random groups of schools and the subsequent assignment of groups to the half-samples $T = 30$ times (Rao & Shao, 1996)—yielding 240 sets of replicates for each sample.

Up to this point, a NEPS data user could have created replication weights for the BRR him-/herself by using the data published, without relying on support from the NEPS methods group. However, the replication weights derived so far have not been corrected for unit nonresponse. In NEPS, because of legal data security regulations, data on nonresponse among schools and students are highly confidential. Studies conducted by the NEPS methods group have revealed that in the sample of students in Grade 5 and 9 school and student nonresponse is systematic (Steinhauer et al., 2013, forthcoming; Aßmann et al., 2012). Neglecting this circumstance when analyzing NEPS data might lead to severe bias in survey estimates. To nevertheless allow for the application of the BRR method, the NEPS method group provides replication weights

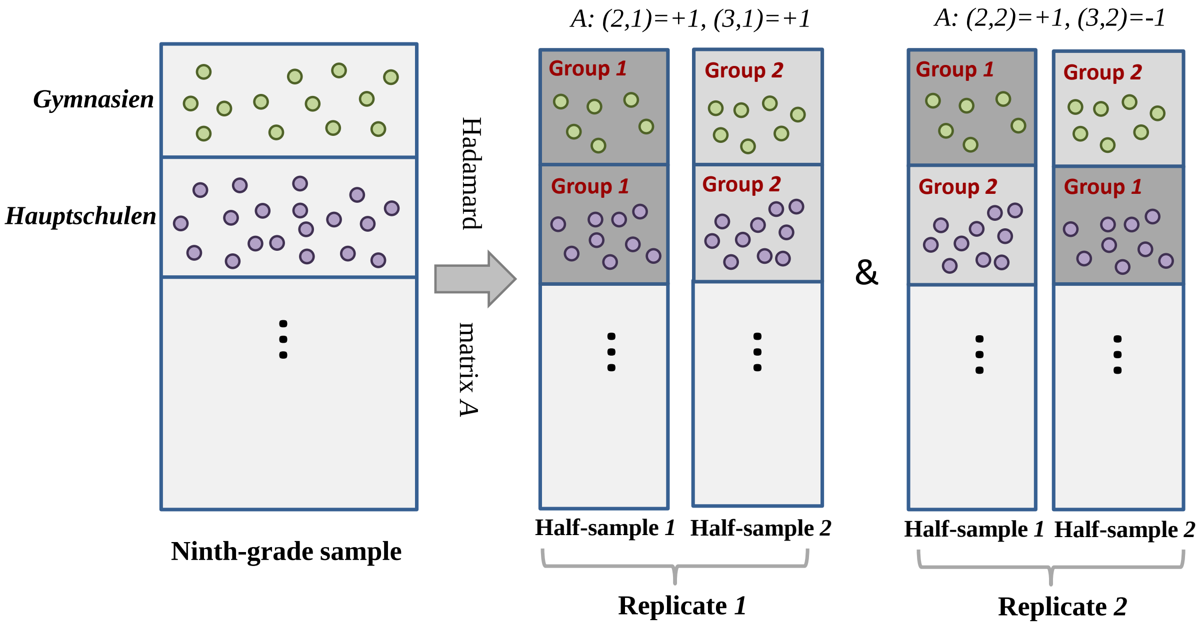

Figure 2: The schools in the strata *Gymnasien* and *Hauptschulen* of the ninth-grade sample are divided into two random groups: Group 1 and Group 2. According to the entries of the Hadamard matrix A these groups are assigned to the half-samples of the replicates.

that have been adjusted to institutional and individual nonresponse. For the nonresponse adjustment of the replication weights we employed the same methods and models as were applied to the sampling weights: We used cell weighting to adjust for nonresponse on the school level and response propensity modeling to correct for nonresponse on the individual level. Both approaches are described in great detail in Steinhauer et al. (2013, forthcoming).

For the construction of half-samples we have applied a BRR variant that uses a perturbation term (see equation (5)), that is, in the distinct half-samples of the replicates schools are weighted differently. Thus, to ensure a reasonable weight adjustment, each school has to enter the nonresponse model weighted accordingly. The corresponding weighting factors $k_{hi}^{(r)}(\epsilon)$, $h = 1, \ldots, H; i = 1, \ldots, n_h; r = 1, \ldots, R)$, can easily be derived from equation (5):

$$k_{hi}^{(r)}(\epsilon) = \begin{cases} 1 + \epsilon\sqrt{\frac{n_h - g_h^1}{g_h^1}}, & \text{if entry } (h, r) \text{ of Hadamard matrix } A \text{ is } +1, \\ 1 - \epsilon\sqrt{\frac{g_h^1}{n_h - g_h^1}}, & \text{if entry } (h, r) \text{ of Hadamard matrix } A \text{ is } -1. \end{cases}$$

For the fifth- and ninth-grade sample any poststratification to external population distributions data was not deemed necessary (Steinhauer et al., 2013, forthcoming). Therefore, the BRR replication weights were not subject to poststratification either.

5 Summary of Weights

For SC3 and SC4 the NEPS provides replication weights for students participating in the first wave of SC3 and for students participating in the first and second wave of SC4. In addition, replication weights are provided for the subset of these students for whom an additional interview with one parent is available. All kinds of weights are given in a trimmed and standardized form.[5] The following table lists the types of replication weights provided within the SUF release Version 1.1.0:

Replication Weights for	Starting Cohort	Label
Students participating in Wave 1	SC3	w1rep_001, ...,w1rep_240
For students participating in Wave 1 & realized parent interview	SC3	w2rep_001, ...,w2rep_240
For students participating in Wave 1 & 2	SC4	w1rep_001, ...,w1rep_240
For students participating in Wave 1 and 2 & realized parent interview	SC4	w2rep_001, ...,w2rep_240

[5] The weights are trimmed and standardized in the same way as the sampling weights of schools and students as part of SC3 and SC4. More details on the respective methods are given in Aßmann et al. (2012).

References

Aßmann, C., Steinhauer, H. W., Kiesl, H., Koch, S., Schönberger, B., Müller-Kuller, A., ... Blossfeld, H.-P. (2011). Sampling design 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* (Vol. 14, p. 51-65). Wiebaden: VS Verlag f"ur Sozialwissenschaften.

Aßmann, C., Steinhauer, H. W., & Zinn, S. (2012, October). *Starting Cohorts 3 and 4: 5th/9th Grade (SC3/SC4), SUF Version 1.0.0, Data Manual [Supplement]: Weighting* (Tech. Rep.). Bamberg: National Educational Panel Study (NEPS). Retrieved from `https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC3/1-0-0/SC3_SC4_1-0-0_Weighting_EN.pdf`

Judkins, D. (1990). Fay's method for varivari estimation. *Journal of Official Statistics*, *6*(6), 223-239.

Kalton, G. (1979). Ultimate cluster sampling. *Journal of The Royal Statistical Society*, *142*, 210-222.

Lee, E., & Forthofer, R. (2006). *Asampling complex survey data* (2nd ed.; L. Shaw, K. Wong, M. Birdsall, & A. Sobczak, Eds.). SAGE Publications.

McCarthy. (1969). Pseudo-replication: half samples. *Review of the International Statistical Institute*, *37*, 239-264.

Rao, J., & Shao, J. (1996). On balanced half-sample variance estimation in stratified random sampling. *Journal of the American Statistical Association*, *91*, 343-348.

Rao, J., & Shao, J. (1999). Modified balanced repeated replication for complex survey data. *Biometrika*, *86*, 403-415.

Rust, K., & Rao, J. (1996). Variance estimation for complex surveys using replication techniques. *Statistical Methods in Med*, *5*, 283-310.

Saigo, H., Shao, J., & Sitter, R. (2001). A repeated half-sample bootstrap and balanced repeated replications for randomly imputed data. *Survey Methodology*, *27*(2), 189-196.

Shao, J. (1996). Resampling methods in sample surveys. *Statistics: A Journal of Theoretical and Applied Statistics*, *27*(3-4), 203-237. (Invited discussion paper)

Shao, J., & Chen, Y. (1999). Approximate balanced half sample and related replication methods for imputed survey data. *The Indian Journal of Statistics*, *61*(1), 187-201.

Shao, J., Chen, Y., & Chen, Y. (1998). Balanced repeated replication for stratified multistage survey data under imputation. *Journal of the American Statistical Association*, *93*(442), 819-831.

Steinhauer, H. W., Aßmann, C., Zinn, S., & Goßmann, S. (2013, forthcoming). Sampling and weighting panel cohorts in institutional contexts. In *Methodological Issues of Longitudinal Surveys.*

Valliant, R. (1987). Some prediction properties of balanced half-sample variance estimators in single-stage sampling. *Journal of the Royal Statistical Society*, *49*, 68-81.

Wolter, K. (2007). *Introduction to variance estimation* (2nd ed.; S. Fienberg & W. van der Linden, Eds.). Springer.

Zinn, S. (2013, forthcoming). Variance estimation with balanced repeated replication: An application to the fifth and ninth grader cohort samples of the National Educational Panel Study. In *Methodological Issues of Longitudinal Surveys*.