

## 4 Sampling designs of the National Educational Panel Study: challenges and solutions

**Christian Abmann · Hans Walter Steinhauer · Hans Kiesel · Solange Koch · Benno Schönberger · André Müller-Kuller · Götz Rohwer · Susanne Rässler · Hans-Peter Blossfeld**

**Abstract:** Research in the German National Educational Panel Study intends to provide the empirical basis for a longitudinal analysis of individuals' educational careers and competencies and how they unfold over the life course in relation to formal as well as nonformal/informal learning environments. In order to track educational developments and decisions over the life span, six main samples serve as the foundation for the characterization and analysis of educational processes. These six main samples include newborns, Kindergarten children, secondary school children (fifth and ninth grade), first-year undergraduate students and adults. They are accompanied

---

© VS Verlag für Sozialwissenschaften 2011

Dr. C. Abmann (✉) · Dipl.-Volksw. H. W. Steinhauer · Dipl.-Pol. S. Koch  
Dipl.-Soz. B. Schönberger · Dipl.-Soz. A. Müller-Kuller  
National Educational Panel Study, University of Bamberg, 96045 Bamberg, Germany  
e-mail: christian.assmann@uni-bamberg.de

Dipl.-Volksw. H. W. Steinhauer  
e-mail: hans-walter.steinhauer@uni-bamberg.de

Dipl.-Pol. S. Koch  
e-mail: solange.koch@uni-bamberg.de

Dipl.-Soz. B. Schönberger  
e-mail: benno.schoenberger@uni-bamberg.de

Dipl.-Soz. A. Müller-Kuller  
e-mail: andre.mueller-kuller@uni-bamberg.de

Prof. Dr. H. Kiesel  
Department of Computer Sciences and Mathematics, Regensburg University of Applied Sciences,  
93025 Regensburg, Germany  
e-mail: Hans.Kiesel@hs-regensburg.de

Prof. Dr. G. Rohwer  
Chair of Sociology, University of Bochum, 44780 Bochum, Germany  
e-mail: Goetz.Rohwer@rub.de

Prof. Dr. S. Rässler  
Chair of Statistics and Econometrics, University of Bamberg, 96045 Bamberg, Germany  
e-mail: susanne.raessler@uni-bamberg.de

Prof. Dr. Dr. h.c. H.-P. Blossfeld  
Chair of Sociology I, University of Bamberg, 96045 Bamberg, Germany  
e-mail: soziologie1@uni-bamberg.de

by several additional samples allowing an analysis of special groups, e.g., special needs pupils. Given that for several of the starting cohorts access to the target population is gained via educational institutions like Kindergartens and schools, multistage sampling approaches reflecting the multistage access to the target populations are implemented. Samples in individual contexts as for the cohorts of adults and newborns are established via register-based stratified cluster approaches. The designs of the implemented sampling strategies are shortly reviewed for each established sample.

**Keywords:** Panel study · Multistage sampling · Indirect sampling · Explicit and implicit stratification · Proportional to size sampling

## **Strategien der Stichprobenziehung im Rahmen des Nationalen Bildungspanels: Anforderungen und Lösungsansätze**

**Zusammenfassung:** Die Forschungsintention des Nationalen Bildungspanels zielt auf die Bereitstellung der empirischen Grundlagen für eine fundierte längsschnittliche Analyse individueller Bildungsverläufe ab. Insbesondere steht die Abbildung individueller Kompetenzentwicklungen, unter besonderer Berücksichtigung des persönlichen und familiären Hintergrundes, sowie der Beziehung zu formalen Bildungsinstitutionen im Vordergrund. Zur Charakterisierung und Analyse von Bildungsentwicklungen und -entscheidungen über den gesamten Bildungsverlauf werden sechs Hauptstichproben etabliert. Diese Hauptstichproben werden jeweils als Zufallsstichproben aus den Kohorten der Neugeborenen, Kindern in Kindergärten, Kindern in der Sekundarstufe an allgemeinbildenden Schulen (fünfte und neunte Jahrgangsstufe), Erstsemesterstudierenden und Erwachsenen gebildet. Eine Ergänzung dieser sechs Hauptstichproben findet durch einige Zusatzstichproben statt, welche eine Analyse spezieller Gruppen erlauben, z.B. Kinder an Förderschulen. Da für etliche dieser Startkohorten des Nationalen Bildungspanels der Zugang zu den Zielpopulationen im institutionellen Kontext von Kindergärten und Schulen erfolgt, beruhen die entsprechenden Stichprobenziehungsverfahren auf einer mehrstufigen Zufallsauswahl, durch welche die institutionellen Gegebenheiten beim Feldzugang berücksichtigt werden können. Für Stichproben, bei denen die Erhebungen im individuellen Kontext stattfinden, erfolgt die Zufallsauswahl im Rahmen von registergestützten Klumpenstichproben. Die zur Anwendung kommenden Stichprobenziehungsverfahren werden kurz für alle zu etablierenden Stichproben dargestellt.

**Schlüsselwörter:** Panelstudie · Mehrstufige Zufallsauswahl · Indirekte Zufallsauswahl · Explizite und implizite Stratifizierung · Größenproportionale Zufallsauswahl

### **4.1 Introduction**

In order to provide an empirical basis for paradigms and theories discussed within life-course research, the National Educational Panel Study (NEPS) establishes six main samples including newborns, Kindergarten children, secondary school children (fifth and ninth grade), students and adults to implement the multicohort sequence design, see Chap. 1 in this volume. Whereas the samples of children in Kindergartens, children in secondary schools, and students at universities cover important episodes in the German educational system, the sample of newborns allows analyzing educational developments before they enter formal educational institutions. Implementation of the multicohort sequence approach designed to span all important episodes in educational life is then

completed by a sample of adults who have already left the formal educational institutions. Although these six main samples are designed to ensure that every individual in the target populations has a chance to be part of the samples, several small but interesting groups will be covered with sample sizes that are too small for statistical analysis with sufficient precision. Hence, several additional samples focus on highlighting these groups of special interest. Also an oversampling of pupils having a high probability to enter vocational tracks after ninth or tenth grade is incorporated. This oversampling aims at sufficient sample sizes for empirical analysis of this heterogeneous group. Surveying the six main cohorts and the additional samples starts in time periods allowing a neat investigation and recording of important educational transitions and decisions from the very beginning of the panel study. In particular, these educational decisions and transitions are entering Kindergarten, the enrollment in elementary school, the transition to secondary schools and vocational tracks, and the transition to higher tertiary education.

Because individual competence tests planned in all studies are costly and time-consuming, samples are based as far as possible on clusters of subjects. It is easier and more cost-efficient to conduct competence tests in larger organizational units such as Kindergarten groups or school classes. In addition, such clusters also allow multilevel analyses of educational processes. Thus, whenever possible, an institutionally based random sample using clusters of Kindergartens, schools, or fields of study at universities is preferred. The cohorts starting in the institutional context of Kindergartens, schools and universities will hence be based on a multistage sampling approach using as sampling units the Kindergarten, the school class or the field of study respectively. In subsequent years, all respondents will be followed up as far as possible in their institutional contexts. Participants leaving the institutional context prevailing at the beginning of the survey will be followed up individually or in their new institutional context. Examples are children entering elementary school earlier or later than average school enrollment, students repeating a grade or students changing their school type during the lower secondary level. All these students will be surveyed yearly, but tested at larger intervals only. However, from a sampling and longitudinal weighting perspective individuals will always be related to their starting cohort and sample. To trace the later educational experiences and careers of persons who are not yet in or have left the educational system already, sampling of individuals will be register-based.

This chapter proceeds as follows. In Sect. 4.2 detailed definitions of the populations and intended sample sizes underlying the main and additional samples are provided. Thereafter, Sect. 4.3 presents some methodological background on the applied sampling strategies including their theoretical properties. Detailed information on the implemented schemes for the main and the additional samples is discussed in Sect. 4.4 and Sect. 4.5 concludes.

## **4.2 Definition of target populations**

### **4.2.1 Main samples**

#### *4.2.1.1 Newborns*

The target population of newborns covers all children who will be born in Germany in the time period from March to August 2012. The plan is to conduct individual testing with

approximately 3,000 children and interviews with their mothers. The first panel wave starts, when the children are six to eight months old.

#### *4.2.1.2 Kindergarten and elementary school children*

Target persons are children at about the age of four years attending Kindergartens in school year 2010/2011 in Germany who are expected to begin schooling in the school year 2012/2013. Because the time period covered by school years varies across federal states with the beginning of the school year ranging from August to September, the age of surveyed children differs across federal states. Handicapped and disabled children generally belong to the target population. However, if their degree of disability makes it impossible for these children to participate in the standardized testing procedure, they may be excluded by the test supervisor. Kindergartens are defined as institutions where children are taken care of for the whole day or a part of it. In these institutions, children are being cared and educated on a regular basis, having full- or part-time working staff with an operating license according to § 45, Volume VIII of the German Social Code Book or an equivalent license. Special needs day-care centers are not considered. Kindergartens are funded by public, non-governmental or private bodies. Note that children receiving day-care regularized by § 43, Volume VIII of the German Social Code Book (Tagesmütter) are excluded. The plan is to survey approximately 3,000 children, their parents and the staff of the Kindergartens involved.

The elementary school children population is defined as all children attending first grade in state-approved or officially recognized elementary schools in school year 2012/2013. It is planned to conduct a survey of approximately 6,000 children and their parents, while gaining information on the school context from their teachers and principals.

#### *4.2.1.3 Secondary school children—fifth and ninth grade*

Target persons are all children attending secondary school in fifth or ninth grade in Germany in the 2010/2011 school year. 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 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. Overall, the plan is to survey approximately 6,800 children in grade five and 13,500 children in grade nine. The number of ninth grade students is higher, because it has to be sufficiently large to permit following up students switching to a vocational track after grade nine with a sample size large enough to allow precise empirical analysis. The survey of children is then complemented by parental interviews and information provided by teachers and principals.

#### *4.2.1.4 First-year undergraduate students*

The population of first-year students is composed of all students (German and non-German) enrolled for the first time in officially recognized and state-approved institutions of

higher education in Germany aiming towards a bachelor's degree, a state examination (Staatsexamen) in medicine, law studies, pharmacy and teaching, or a diploma or master's degree in Catholic or Evangelical theology in the 2010/2011 academic year. A special focus is set on students with a non-traditional admission certificate, see Chap. 17 in this volume for more details. Students attending universities, universities of technology, or universities of applied sciences run by federal ministries or federal states for members of their public services are excluded. The plan is to establish a sample of approximately 16,500.

#### 4.2.1.5 Adults

The target population consists of all people living in private households in Germany who belong to the birth cohorts of the years 1944 to 1986. The plan is to survey approximately 300 people for each birth year. Access to this target population is gained via three channels. The first channel is the already existing sample in the survey "Arbeiten und Lernen im Wandel" (ALWA) conducted by the Institute for Employment Research (IAB) in 2009, see Antoni et al. (2010). This survey covers the birth cohorts ranging from 1955 to 1986. In the first survey wave of the adults' cohort this core sample is being refreshed and for the early birth cohorts of 1944 to 1954 an additional (augmentation) sample is drawn. In summary approximately 13,000 adults shall be surveyed.

### 4.2.2 Additional samples

#### 4.2.2.1 Special needs children

The survey of school children is completed via consideration of an additional sample of children attending special schools. The focus is particularly put on special needs tracks handling learning disabilities. Access to this population is gained via special schools with federal-state-specific provisions explicitly for children with learning disabilities. Overall, 80% of children attending special schools have a diagnosed learning disability, therefore constituting the largest group of children in these schools. The plan is to survey approximately 1,500 children in grade nine and 750 children in grade five in conjunction with their parents and teachers in school year 2010/2011. In the school year 2012/2013 also first grade students shall be surveyed.

#### 4.2.2.2 School reform studies

Two additional samples are focusing on the effect of structural school reforms. On the one hand, the effect of curricular reforms in Thuringia taking place in 2010 and 2011 shall be studied. The target population is defined to contain all Gymnasium students attending the twelfth grade either in 2009/2010 (last cohort before reforms were implemented) or in 2010/2011 (first cohort after implementation of reforms). On the other hand, the effect of a school reform in Baden-Wuerttemberg will be analyzed, where the Gymnasium will be shortened from nine to eight years. The affected population consists of all pupils attending the twelfth or thirteenth grade in Gymnasien in 2011/2012. Furthermore, all students

attending the thirteenth grade in 2010/2011 or the twelfth grade in 2012/2013 are defined as control group populations.

### 4.3 Methodological background

#### 4.3.1 Stratified multistage sampling based on explicit and implicit stratification

Probability sampling is an essential prerequisite for conducting educational surveys. A sampling design with sample size  $n$  is called simple random sampling, if every combination of  $n$  units from the population is equally likely to be chosen. In most cases it is neither possible nor desirable to draw a simple random sample. Since the main goal of any survey is to come up with estimates as precise as possible, sampling designs that differ from simple random sampling might increase the precision of estimates. On the other hand, simple random sampling might be infeasible when no sampling frame including all population units is available. In the context of sampling school children, simple random sampling is unfeasible, since no listing of all school children is available, and it is also undesirable, since other designs (e.g., stratified sampling) are more efficient.

Stratified sampling means that the target population is divided into subpopulations (called strata), and then samples are drawn independently from each stratum. There are several possible strategies to allocate the total sample size to the different strata. If the sample sizes are proportional to the population sizes within each stratum, this is called proportional allocation and results in equal sampling rates across strata (and usually in more precise estimates for characteristics of the whole population). Since the precision of estimates for subpopulations mainly depends on the sample size within these subpopulations, varying sampling rates across strata might be reasonable to increase the precision of estimates for special groups. (Using higher sampling rates among certain strata is called oversampling). Thus, stratified sampling has two main advantages: it generally leads to more precise estimates of the whole population, and it allows using different sampling rates in different strata to control precision of sub-group estimates.

In cluster sampling, a sample (possibly stratified) is drawn in such a way that clusters of units are selected at once. In one-stage cluster sampling, every unit in each selected cluster is surveyed. Usually, however, only a sample of units within each selected cluster will be drawn; this is called two-stage sampling or (if sub-clusters are drawn within the first stage clusters) multistage sampling. Cluster sampling is likely to reduce the precision of estimates compared to simple random sampling (with equal sample size), since units within clusters tend to be more similar than units in different clusters. This is often outweighed, however, by the advantages of cluster sampling: first, in some applications lists of clusters are the only available sampling frame (for surveys of school children, lists of schools are available but complete lists of all children are not); second, surveying of clusters is usually much cheaper than surveying of individual based samples (e.g., testing children within the same school can be done faster and cheaper than testing the same number of children in different schools); third, cluster sampling is appropriate to measure context effects. For all three reasons, the NEPS uses multistage cluster sampling for most of its samples.

Stratification and cluster sampling may be combined, resulting in stratified multistage cluster samples. Within each stratum, units or clusters might be sampled with designs different from simple random sampling. Often, the sampling frame is ordered by some variables, and then a systematic sample is drawn. The precision of the resulting estimates is similar to the result of stratification with proportional allocation, and therefore this procedure is called implicit stratification in contrast to explicit stratification as described above. (Usually implicit stratification is done within explicit strata.)

Unbiased estimation of population characteristics based on a sample is only possible if every unit in the target population has a non-zero probability to be part of the sample (these probabilities are called inclusion probabilities). The inclusion probabilities depend on the way the sample is being selected. In simple random sampling and also in systematic sampling, every unit in the sample has the same inclusion probability; the same is true for stratified simple random sampling with proportional allocation of sample size. Different sampling rates among strata result in unequal inclusion probabilities. In multistage cluster sampling, first stage clusters are often selected with inclusion probabilities that are proportional to some measure of cluster size; this is called pps (“probability proportional to size”) sampling. For a thorough treatment of the theory and the pros and cons of different sampling designs, see e.g., Cochran (1977); Lohr (2009) or Särndal et al. (2003).

The use of stratified multistage approaches with pps sampling is common for educational surveys given the nested and hierarchical structure of educational systems. For example, within the Trends in International Mathematics and Science Study (TIMSS) schools were drawn with probability proportional to student enrollment at the first stage. At the second stage classes (as clusters of students) were then selected with equal probability, see Joncas (2008). For the Programme of International Student Assessment (PISA) schools were drawn by pps sampling, where the measure of size is related to the number of target pupils in each school (OECD 2009). A further example is the National Educational Longitudinal Survey of 1988 (NELS:88), where a two-stage design with schools at the first stage and students within schools at the second stage was applied. The inclusion probabilities for the first stage units were proportional to their estimated eighth grade enrollment, and students were selected with equal probabilities in the second stage (Thurgood et al. 2003). All of these studies used explicit stratification; the TIMSS and PISA studies also adopted implicit stratification.

Regardless of the chosen sampling design (given positive inclusion probabilities for every population unit), the Horvitz-Thompson estimator calculates an unbiased estimate for the population total  $t_y$  of any variable  $Y$ , using the sample values  $y_1, \dots, y_n$ :

$$\hat{t}_{Y,HT} = \sum_{i=1}^n \frac{y_i}{\pi_i} = \sum_{i=1}^n w_i y_i,$$

where  $\pi_i$  is the inclusion probability of sample unit  $i$  (in multistage samples,  $\pi_i$  is the product of the conditional inclusion probabilities at each stage); the inverse of  $\pi_i$  is called the sampling or design weight  $w_i$  of unit  $i$  (the intuition behind this is that sample unit  $i$  “represents”  $w_i$  elements of the population). Thus, the estimate for the population total is the sum of the sample values weighted by the sampling weights. Using the sampling weights, estimates for other population characteristics (e.g., means, quantiles, correlations) may be constructed as well. There are also more advanced estimators taking

non-response into account or incorporating auxiliary information (see e.g., Särndal et al. 2003). It is also important to note that the variance of the Horvitz-Thompson estimator (and thus the precision of the estimates) does not only depend on the  $\pi_p$  but also on the specific sampling design.

#### 4.3.2 Indirect sampling

In general, sampling designs as described in the previous section are based on a complete list (sampling frame) of all units (or clusters of units) representing the target population. For different reasons it may happen that a sampling frame is not available, thus hindering a direct application of established sampling strategies. In this situation Lavallée (2007) suggests using a different approach called indirect sampling. In the NEPS, indirect sampling is used for establishing the sample of Kindergarten children.

Suppose there exists a population  $U_A$  with an available sampling frame, and this population is somehow “linked” to the target population  $U_B$ . A natural idea is then to draw a sample  $s_A$  from population  $U_A$  and subsequently choose all elements from  $U_B$  that are linked to elements in  $s_A$  and define them as the sample  $s_B$  from  $U_B$ . Although the calculation of inclusion probabilities for units in  $s_B$  is usually difficult or unfeasible under this setting, it is nevertheless possible to construct an unbiased estimator for population totals of  $U_B$ .

To formally describe the connection between the two populations, let  $\theta_{ab} \geq 0$  represent the link between  $a \in U_A$  and  $b \in U_B$  (If  $\theta_{ab} = 0$ , there is no link existing between  $a$  and  $b$ ). In our application,  $U_B$  is the population of Kindergartens,  $U_A$  is the population of elementary schools,  $\theta_{ab}$  may be defined as the number of children moving from Kindergarten  $b$  to elementary school  $a$  in some reference year. In general, how best to define the links  $\theta_{ab}$  depends on the application at hand.

For every  $b \in U_B$ , let  $\theta_{+b} := \sum_{a \in U_A} \theta_{ab}$  be the sum of all links from  $U_A$  to  $b$ . We assume that  $\theta_{+b} > 0$  for all  $b \in U_B$ , i.e. there exists a link to every  $b \in U_B$ ; otherwise unbiased estimation for  $U_B$  is impossible.

The key observation is that the total of any variable  $Y$  in population  $U_B$  might be written as follows:

$$t_y = \sum_{b \in U_B} y_b = \sum_{b \in U_B} \left( y_b \cdot \underbrace{\sum_{a \in U_A} \frac{\theta_{ab}}{\theta_{+b}}}_{=1} \right) = \sum_{a \in U_A} \sum_{b \in U_B} \frac{\theta_{ab}}{\theta_{+b}} y_b = \sum_{a \in U_A} \tilde{y}_a = t_{\tilde{y}},$$

$$\text{with } \tilde{y}_a := \sum_{b \in U_B} \frac{\theta_{ab}}{\theta_{+b}} y_b.$$

Thus, the total of some variable  $Y$  in population  $U_B$  can be written as the total of the corresponding variable  $\tilde{y}$  in population  $U_A$ . Since inclusion probabilities for  $s_A$  are known, the Horvitz-Thompson estimator may be used for unbiased estimation of  $t_{\tilde{y}}$  and thus also of  $t_y$ . The (unbiased) indirect sampling estimator for the total of  $Y$  in  $U_B$  is defined as



$$\hat{t}_{Y,IS} := \hat{t}_{\bar{y},HT} = \sum_{a \in s_A} \frac{\bar{y}_a}{\pi_a} = \sum_{a \in s_A} \sum_{b \in U_B} \frac{\theta_{ab}}{\theta_{+b}} \frac{y_b}{\pi_a} = \sum_{b \in s_B} w_{b_s} y_b$$

with weights  $w_{b_s} = \sum_{a \in s_A} \frac{\theta_{ab}}{\pi_a \cdot \theta_{+b}}$  for  $b \in s_B$ . These weights are sample dependent and in general different from the inverse inclusion probabilities (although they are equal in expectation).

If the elements of  $U_B$  are actually clusters of individuals, all units within cluster  $b$  are assigned the same sampling weight  $w_{b_s}$ . It is also possible to add another sub-sampling stage within the clusters in  $s_b$  and then change the sampling weights accordingly (two-stage indirect sampling; see Lavallée 2007 for details). Indirect sampling has been applied in different surveys, e.g., in the Survey of Labour and Income Dynamics (SLID) and the Project to Improve Provincial Economic Statistic (PIPES), both conducted by Statistics Canada; however, the NEPS is the first application of indirect sampling techniques in the context of educational surveys.

## 4.4 Sampling strategies

### 4.4.1 Main samples

#### 4.4.1.1 Newborns

The sampling of newborns will be conducted as a stratified cluster sampling. Clusters are defined as municipalities, where sampling is stratified according to regional classification criteria, such as federal states and counties. Sampled municipalities are then asked to provide the register-based address information on the defined target population. The plan is to conduct the survey of newborns in approximately 120 municipalities, whereby this number may be subject to change due to demographic developments.

#### 4.4.1.2 Kindergarten and elementary school children

The sampling of Kindergarten and elementary school children is designed to provide random samples for each of the two underlying populations. Surveying of elementary school children will start in 2012, when most of the sampled Kindergarten children will move to elementary schools. The design of the two samples is therefore coordinated in such a way that a large fraction of the Kindergarten sample will show up in the elementary school sample. Such an overlap yielding institutional context information for school children already surveyed in Kindergartens will extend the range of research questions to be answered by longitudinal analysis.

Since Kindergartens and elementary schools are connected in form of children passing over from Kindergartens to elementary schools, an indirect sampling approach using this link information is defined for establishing both samples. Thereby the first sampling stage for establishing a sample of Kindergarten children simultaneously provides a sample of schools allowing access to elementary school children. Given that the survey of Kindergarten children starts two years earlier, the elementary school children sample increases

the total number of available cases of the then approximately six year old Kindergarten children. The measure of size for pps sampling of elementary schools is defined as the number of children attending first grade based on the available frame information referring to school year 2008/2009. Furthermore sampling is based on implicit stratification of schools according to federal states, regional classification and organizing institution. All children attending first grade in the sampled schools are surveyed. Since schools serve as hinges to get access to Kindergartens via indirect sampling, the defined measure of size increases the probability that each school is linked to at least one Kindergarten providing thus access to the population of Kindergarten children.

Based on the sample of elementary schools as the first sampling stage of the implemented indirect sampling approach for surveying Kindergarten children, the second stage is administered as follows. Sampled schools are asked to list all Kindergarten institutions from which children have entered these schools in school year 2009/2010. From these lists of Kindergartens, sampling Kindergartens is performed proportional to the number of trespassing children. Based on a small scale simulation study referring to data on Bavarian Kindergartens, proportional sampling has been chosen in order to enlarge the rate of children surveyed in the context of Kindergartens and schools extending thereby the range of possibilities for longitudinal analysis. Note that this procedure only provides access to the population of Kindergarten children in institutions already established in 2009/2010. Given the selection probabilities for schools at the first stage and Kindergartens at the second stage, the resulting two stage indirect sampling weights are then additionally based on the number of trespassing children between Kindergartens and elementary schools and the total number of children per Kindergarten entering schools in the 2009/2010 school year.

#### *4.4.1.3 Secondary school children—fifth and ninth grade*

Sampling children within an institutional context provides important information on the institutional background, thereby increasing the range of questions to be analyzed with the survey. Furthermore, as an important aspect for a longitudinal survey, tracking of children is easier and hence combined with a reduction of administrative survey costs.

The large variety of federal-state-specific school systems is a challenge for sampling fifth and ninth grade children. Since educational policy is the responsibility of each single federal state, many different school types related to different transitions between elementary and secondary schooling institutions make up the set of institutions providing access to the target population of fifth and ninth grade students. For the purpose of sampling, the population of schools is therefore stratified by school type. Six school type strata are defined in total. The first stratum comprises all Gymnasien, the second stratum is set up by all Hauptschulen, the third stratum refers to all Realschulen, the fourth to Gesamtschulen, the fifth includes all schools offering all tracks of secondary education except the academic track (Schulen mit mehreren Bildungsgängen). Finally, the sixth explicit stratum comprises all schools providing schooling to fifth grade students, but not to ninth grade students. In addition to explicit stratification according to school types, an implicit stratification based on the same variables as for sampling elementary schools, i.e. fed-

eral states, regional classification and organizing institution, is adapted for sampling secondary schools. The definition of these six explicit strata allows serving two important aspects. The first is related to establish the sample of ninth grade children as the starting point of a longitudinal survey of young adults entering vocational tracks 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 and Gesamtschulen an oversampling of children attending these school types is incorporated. The second aspect served via the explicit stratification is the possibility to reach fifth and ninth grade children population via the same set of schools reducing administrative survey costs for schools establishing the access point to the longitudinal study of the population of fifth grade children, but also representing all children attending fifth grade classes in Germany. Given the first stage sample of schools, at the second stage two school classes within each school are sampled, if at least three classes are present, otherwise all classes are surveyed.

A measure of size for sampling schools is defined, which yields a variation in design weights within strata as low as possible. Given the sampling of two classes within schools a reasonable measure of size for sampling schools is proportional to  $\frac{N_{classes,i}}{\min(N_{classes,i};2)}$  ( $N_{classes,i}$  denotes the number of classes in school  $i$ ), which results in combination with a simple random sampling of (at most) two classes within schools in roughly equal design weights within each stratum. Unfortunately, some variance of the weights cannot be avoided completely. Since sampling of schools is based on a frame provided by the statistical offices of the federal states referring to information available in school year 2008/2009, the current situation in schools is not mirrored in the available school information to the full extent. For some selected schools the number of classes listed for the seventh grade differs from the number of ninth grade classes actually surveyed in school year 2010/2011 due to fluctuations in student enrollment. Using available frame information on the development of the number of classes referring to a particular schooling cohort, a measure of size proportional to the number of classes in the seventh grade in school year 2008/2009 showed within a simulation study substantially less variation in design weights for pupils than alternative measures of size based on other characteristics like average class size. Also currently implemented school reforms often causing a change of the school type are not fully reflected within the available school frame. Given a change of school type occurs for a sampled school, but classes in fifth or ninth grade are still available, participation of the school in the survey is continued.

A further issue is handling institutional refusals to participate within the survey. Schools possibly refuse to participate in order to avoid further workload arising from participating in other studies as well. Since the resulting sample size reduction on the level of students should be compensated in order to ensure sample sizes large enough for analyzing subpopulation differences, replacement schools are defined in advance according to a matching rule. The matching rule defines replacement schools as schools from the same explicit and similar implicit stratum as the original school. Although the use of replacement schools does not eliminate the risk of bias due to non-response, employing implicit and explicit stratification increases the chances that any school's replacement would have similar characteristics.

#### 4.4.1.4 First-year undergraduate students

Sampling first-year undergraduate students with ordinary admission certificates is achieved via a stratified cluster sampling approach. Note that next to these students, also all first-year students with a non-traditional admission certificate shall be surveyed based on postal recruitment, see Chap. 17 in this volume for details. For sampling purposes, a cluster is defined as a field of study to be attended at a higher educational institution. Within each cluster all students are to be surveyed. In order to achieve high response rates, sampled students are approached by two different contact modes: first, all students are contacted by mail. Second, central first-year courses are attended by field workers to ask for participation. In a pilot study, this twofold recruitment process yielded higher participation rates, as well as a higher panel attendance. The student cohort has been set up to incorporate an oversampling of students attending teaching tracks and students attending private higher educational institutions, i.e. private universities and private universities of applied sciences. This objective is addressed via setting up a first stratification level grouping clusters according to their educational institutions. This first stratification level defines four strata: Stratum  $h_1$  comprises the clusters linked to teaching tracks. Stratum  $h_2$  is set up to include all fields of study attendable via public universities, while stratum  $h_3$  summarizes all fields of study offered via public universities of applied sciences. Finally, stratum  $h_4$  comprises all study tracks offered by private universities or private universities of applied sciences. This level of stratification allows oversampling students in teaching tracks and attending private higher educational institutions using different sampling rates for clusters in the four different strata.

However, given the heterogeneous distribution of students across the officially listed fields of study sampling within the defined strata would result in a large sampling variation concerning the coverage of the range of fields of study within the sample. Hence, a further level of stratification is introduced, where strata are defined by groups of related fields of study. This stratification has been accompanied by an exclusion of clusters with less than 30 inscribed students in 2008/2009. In summary, the 60 officially listed fields of study are grouped in several study groups per first stage stratum. Thereby strata  $s_1$  to  $s_3$  group fields of study in stratum  $h_1$ , strata  $s_4$  to  $s_{19}$  correspond to first stage stratum  $h_2$ ,  $s_{20}$  to  $s_{26}$  comprise fields of study within stratum  $h_2$  and grouping in stratum  $h_4$  is achieved via setup of groups  $s_{27}$  to  $s_{29}$ . In order to reproduce the distribution of students across the fields of study and ensure homogenous inclusion probabilities within strata  $h_1$  to  $h_4$ , an according allocation of the number of clusters to be drawn within each stratum is set up.

This allocation is also necessary to incorporate the planned oversampling. In particular the number of clusters  $c_{h_i}$  sampled within stratum  $h_i$  is calculated according to

$$c_{h_i} = \frac{P_{h_i}}{\frac{1}{k_{h_i}} \sum_{k=1}^{k_{h_i}} N_k},$$

dividing planned sample size  $P_{h_i}$  in stratum  $h_i$  by average cluster size measured in number of first-year students  $N_k$  in 2008/2009 for each cluster. This results in 54 clusters to be sampled for stratum  $h_1$ , and in 73 clusters for stratum  $h_4$ . For strata  $h_2$  and  $h_3$ , where no oversampling is adopted, a total of  $n_{23} = 348$  clusters to be sampled have been found sufficient to generate the planned gross sample sizes, where clusters are allocated

proportionally to the number of overall clusters in both strata resulting in 203 clusters to be sampled in stratum  $h_2$  and 145 clusters in stratum  $h_3$ . For each substrata the number of clusters to be sampled from strata  $h_i, i=1, \dots, 4$  are allocated according to

$$c_{s_j} = c_{h_i} \frac{K_{s_j}}{K_{h_i}},$$

where  $K_{s_j}$  denotes the number of clusters in stratum  $s_j$  and  $K_{h_i}$  the number of clusters in first stage stratum  $h_i$ .

The following strategy is defined to handle institutional non-participation. Since the refusal of a university to participate causes the loss of the fields of study sampled at this specific university, only institutions are eligible for replacement, which allow maintaining the original sample composition with regard to the sampled departments and fields of study. For every combination of sampled fields of study at a particular higher educational institution, all institutions offering the same combination of fields of study within the frame are listed, irrespective of whether the institutions have already been sampled or not. Institutions not sampled are given preferential consideration in the choice of replacements. Given that multiple possible replacement institutions offer the combination of fields of study to be replaced, the replacement institution is defined as the one with the smallest difference in numbers of enrolled students compared to the non-participating institution.

#### 4.4.1.5 Adults

Given the three channels (ALWA, refreshment, augmentation) providing access to the defined target population, for each sample connected to a particular channel a stratified two-stage cluster sampling approach is implemented. The 12,429 German municipalities existing in 2008 are defined as elementary sampling units. Stratification according to federal states and a classification of urbanization (BIK scale) is incorporated. Sampling of elementary sampling units is achieved via sampling of artificial units called sample points. Sample points are defined as regional entities gathering the same number of target persons, where one or more sample points are assigned to each municipality depending on its size. Thus it is possible that multiples of these entities are drawn from single municipalities. This procedure connects multiple sample points of the target population at the level of municipalities. The number of sample points as artificial elementary sampling units to be sampled within each explicit stratum are allocated proportional to the number of target persons. Sampling of persons as secondary sampling units is performed via systematic sampling.

#### 4.4.2 Additional samples

##### 4.4.2.1 Special needs children

Sampling of children in special needs tracks dealing with learning disabilities is performed via cluster sampling of schools, where all children attending the considered fifth and ninth grade are surveyed within sampled schools. Given the small average size of clusters related to the small number of children per class attending on average special

schools, the defined sampling approach has to address the trade-off between sample size and number of sampled schools causing a large fraction of related survey costs. The trade-off is balanced via definition of a measure of size for sampling schools proportional to the squared number of children attending the ninth grade classes. Just like for the main sample schools are used to gain access to the population of ninth and fifth grade children. As a larger sample size is planned to be established concerning ninth grade children in order to ensure sufficient sample size for these children when entering vocational tracks, only a subsample of schools is chosen to provide access to fifth grade children. This subsample of special schools is chosen via simple random sampling, thus establishing a two-stage sampling procedure concerning the sampling of fifth grade special needs children. In addition to the sample of ninth and fifth grade children, a sample of first grade special needs children will be established in 2012 using a similar sampling approach.

#### *4.4.2.2 School reform studies*

For sampling students affected by curricular reforms in Gymnasien in Thuringia in 2010 and 2011 two strata were set up. The first stratum comprises twelfth grade students attending Gymnasien focusing on natural sciences. The second stratum contains all remaining twelfth grade students attending Gymnasien with no explicit natural science focus. Access to these students is gained via simple random sampling of 30 Gymnasien. A sample of Gymnasien students affected by the shortening of the grammar school track in Baden-Wuerttemberg coming into effect in 2012 is established via 50 schools. This set of schools provides access to the four cohorts participating within the study, the thirteenth grade students in 2010/2011, the twelfth and thirteenth grade students in 2011/2012, and the twelfth grade students in 2012/2013. Since for each cohort a gross sample of 40 students shall be sampled in each school, the measure of size for schools is defined proportional to the average number of students enrolled in grades ten, eleven and twelve in school year 2009/2010. The 40 students from each school and cohort are then selected via simple random sampling.

## **4.5 Conclusion**

For establishing random samples of all mentioned populations, where access is gained via educational institutions, multistage and indirect sampling techniques, as well as stratified cluster sampling techniques are applied. These sampling techniques are especially suited to incorporate the institutional settings into the sampling process. As long as individuals are participating in formal educational institutions such as Kindergartens, schools, or institutions of higher education, well-defined clusters exist and are used accordingly. Multistage sampling procedures can also be tailored to cope with the different focuses put forward by the six main samples and the additional samples. The six main samples are established to implement the multicohort sequence design facilitating educational life-course analysis within the defined underlying target populations. Analytical potential is enriched via additional samples of several important groups of interest. In the absence of a complete listing or sampling frame, indirect sampling ensures the availability of a ran-

dom sample. Furthermore, indirect sampling is especially suited to simplify the tracking of the sample of Kindergarten children also within the institutional context of elementary schools. As context information on elementary schooling is provided for some children already surveyed in Kindergartens, indirect sampling is of special value for longitudinal analysis within the NEPS. Stratified cluster sampling strategies are employed to establish a sample of first-year students at tertiary level institutions and a sample for the individual-based survey of adults and newborns. This sampling strategy makes it possible to deal with the splattered landscape of different fields of study not present at all tertiary educational institutions.

## References

- Antoni, M., Drasch, K., Kleinert, C., Matthes, B., Ruland, M., & Trahms, A. (2010). *Arbeiten und Lernen im Wandel, Teil I: Überblick über die Studie* (FDZ-Methodenreport 5/2010). Nürnberg: Forschungsdatenzentrum (FDZ) der Bundesagentur für Arbeit im Institut für Arbeitsmarkt- und Berufsforschung.
- Cochran, W. G. (1977). *Sampling techniques* (3rd ed.). Hoboken: Wiley.
- Joncas, M. (2008). TIMSS 2007 sample design. In J. F. Olson, M. O. Martin, & I. V. Mullis (Eds.), *TIMSS 2007 technical report* (pp. 77–92). Chestnut Hill: TIMSS and PIRLS International Study Center.
- Lavallée, P. (2007). *Indirect sampling*. New York: Springer.
- Lohr, S. (2009). *Sampling: Design and analysis* (2nd ed.). Boston: Brooks/Cole.
- OECD (2009). *PISA 2006 technical report*. Paris: OECD.
- Särndal, C.-E., Swensson, B., & Wretman, J. (2003). *Model assisted survey sampling*. New York: Springer.
- Thurgood, L., Walter, E., Carter, G., Henn, S., Huang, G., Nooter, D., Smith, W., Cash, R. W., & Salvucci, S. (2003). National Education Longitudinal Study of 1988 (NELS:88). In M. Seastrom, T. Phan, & M. Cohen (Eds.), *NCES handbook of survey methods* (pp. 53–66). Washington: U.S. Department of Education, National Center for Education Statistics.