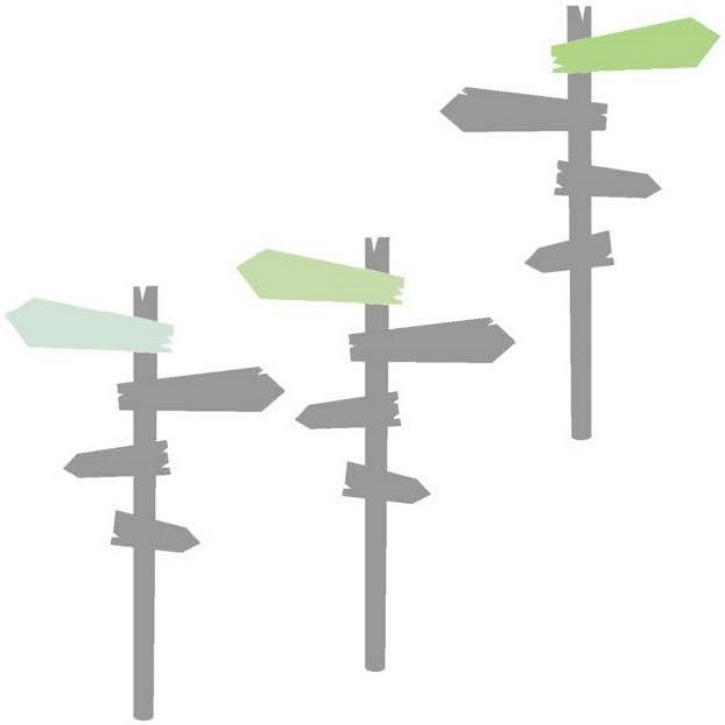
UNIVERSITÄT
BERN

Documentation of the DAB panel study

Bern, 12.2021 (Version 5.0)



Citation:

DAB panel study (2021): Documentation of the DAB panel study (Data-Release Version 5.0). University of Bern, Institut für Erziehungswissenschaft, Abteilung Bildungssoziologie.

Inhaltsverzeichnis

1 Introduction

This document provides information on the data collected in the DAB panel study for all interested in the scientific profile of the DAB panel study and for researchers who are planning to conduct empirical analysis with the scientific-use-files. It serves as a technical documentation and, in addition to the wavespecific codebooks and the variable documentation, as an orientation for empirical analysis.

2 Project overview

2.1 Project goals

Since 2012, the DAB panel study has collected longitudinal data on the vocational and educational situation of adolescence in German-speaking Switzerland. The sampled individuals have been observed since their eighth school year and so far have been surveyed nine times regarding their current training situation as well as their educational and career aspirations. From a life course perspective, the DAB study contributes theory-based empirical evidence for the clarification of relevant questions in the fields of education, labour market and occupational research that previously could not be examined with the available longitudinal data in Switzerland.

The DAB study asks which factors influence the choice of education programmes at the transition from lower secondary education to upper secondary level and potentially tertiary level. The choice of vocational or school based training at the end of compulsory schooling is determined on the one hand by characteristics of the training/labour market and its selection mechanisms, and on the other hand by individual vocational and educational choices. In the first phase of the DAB panel study the interplay of various factors that contribute to the vocational and educational decision (DAB-I) was examined. Individual characteristics of social origin, different competences and school performance, individual interests and preferences as well as askriptive characteristics such as gender or migration background, previous educational decicions and the regional opportunity structures are taken into account.

The continuation of the DAB panel study (DAB-II/III/IV) examines the progress of vocational training or continued general education after the transition to upper secondary level. The pathway of post-compulsory education as well as the successful attainment of educational qualifications at upper secondary level, the continuation of education and training at tertiary level and the transition into employment and the labour market are of particular interest. On the other hand, the data from the follow-up surveys (DAB-II/III/IV) also enable analyses with regard to the change and stability of training plans and career aspirations as well as their realisation. The data collected within the framework of the DAB panel study not only allow a description of the trajectory of post-compulsory educational attainment and entry to employment, but also analyses with regard to so far unclarified questions concerning the structure and contingency of school and vocational training pathways as well as concerning educational returns and decision-making behaviour.

2.2 Theoretical and conceptual background

The structural-individualistic action model in a dynamic multi-level design has proven to be integrative and expandable in theoretical terms for the explanation of educational decisions from a life course perspective (R.

Becker 2012a; R. Becker 2012b). From a structural-individualistic perspective, two sub-processes can be distinguished concerning the realisation of educational and occupational aspirations: firstly, the formation and developement of aspirations and secondly, the process of educational and occupational decision-making (R. Becker 2000; R. Becker 2003). The distinction between the two sub-processes is theoretically appropriate and can be seen as standard in the relevant educational research. The structural-individualistic multi-level model is applied to explain inequalities regarding the entry into employment as well as the decicion to remain in the education system. This theoretical model is enriched with further theoretical approaches of Boudon (1974), Erikson und Jonsson (1996), Breen und Goldthorpe (1997) and Esser (1999). In the DAB panel study, these considerations are applied to educational decisions after compulsory schooling. Theoretical approaches used are: the theory of primary and secondary effects of origin (Boudon 1974), Status Position Theory (Keller und Zavalloni 1964), the Wisconsin model for idealistic and realistic aspirations (Sewell u. a. 1957), Human Capital Theory (G. S. Becker 1964; Mincer 1974), Signal and Filter Theory (Arrow 1973), and the Job Competition Model or Labour Queue Model (Thurow 1975).

While DAB-I focused on the formation and developement of aspirations within the last one and a half years of compulsory schooling, DAB-II investigated their actual realizationation. In DAB-III and -IV, in addition to the completion of post-compulsory training, both the educational and career aspirations prior to completion of initial education at upper secondary level as well as their further implementation are considered and analysed. In this context, it is important to clarify who can realise or must revise their educational or occupational aspirations. Of particular interest are the underlying mechanisms which, as a consequence, contribute to or, at most, reinforce inequalities in educational and occupational opportunities according to social origin, gender and migration background.

2.3 Project structure and funding

The DAB panel study is financed by the State Secretariat for Education, Research and Innovation (SERI) and is institutionally located at the Department of Sociology of Education at the Institute for Educational Science at the University of Bern. So far the DAB panel study covers five funding periods:

DAB-I: 10/2011–09/2013 (BBT-Project-Nr.: BB.2011.0117)

DAB-II: 07/2014–06/2015 (SERI-Project-Nr.: BB.2013.0112)

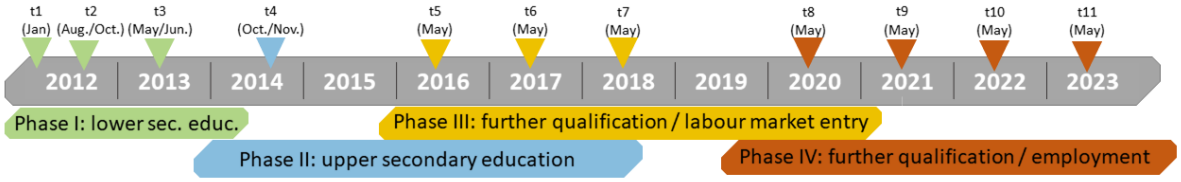
DAB-IIIa: 04/2016–03/2017 (SERI-Project-Nr.: 1315000723)

DAB-IIIb: 04/2017–10/2020 (SERI-Project-Nr.: 1315001039)

DAB-IV: 11/2019–06/2024 (SERI-Project-Nr.: 1315001844)

The overall responsibility for the project lies with Prof. Dr. Rolf Becker as main applicant, supported by the co-applicant Dr. David Glauser. Since autumn 2017, Dr. Sara Möser has been responsible for the project management of the DAB panel study.

2.4 Survey design



The data of the first three waves of the DAB panel study were collected via classroom surveys that were administered as online questionnaires. All students of school classes in a random sample were interviewed. The survey took place in the classroom and was conducted by the corresponding teacher according to a prepared script on the procedure and technical details. The administrative effort of preparing and conducting the surveys was relatively high for teachers, therefore, incentives were sent to the teachers before the first and before the second survey, in order to increase their willingness to participate. In addition to the administration of the online questionnaire in the classroom, the teachers provided access to the grades of respondents. When available the teachers also provided information on test scores of the standardized achievement test «Stellwerk» in German and mathematics. Respondends who changed class or school after the first survey were invited to participate in an online version of the second and third surveys via post.

Additionally the respondents' parents were also interviewed in the first survey. A PAPI questionnaire with a reply envelope and the login details to a online questionnaire were given to the students after the first survey.

With the completion of basic compulsory education after grade 9, the survey design was adapted from classroom to individual survey. From the fourth wave onwards the survey was administered in a sequential mixed-mode design. In a first step, the respondends were first invited via post and e-mail to answer the online questionnaire ("Computer Assisted Web Interviews CAWI). In a second step, those respondents who did not complete the questionnaire within 10-14 days were contacted by telephone. The telephone interview ("Computer Assisted Telephone Interviews CATI) was identical in content to the online questionnaire. As a third survey instrument, a short questionnaire ("paper and pencil interviews PAPI) was sent by post in waves 6 and 7 following the online and CATI survey. Reminder letters were sent out regularly during all three field phases. The respondents were informed by post that they would be contacted by telephone if they did not complete the online questionnaire. They were also reminded to participate via e-mail and, from the sixth survey wave onwards, also by SMS.

With the exception of the fifth wave, for which Limesurvey was used, the online questionnaires were programmed with the survey software UNIPARK from Questback. The CATI surveys of the fourth and seventh wave were carried out by the survey institute MIS-Trend in Bern, the sixth wave by the LINK Institute in Lucerne and the fifth wave was administered by the DAB team in the university's own CATI laboratory.

In the continuation of the DAB panel study (DAB-II/III/IV), material incentives were used in accordance with the logic of the "Tailored Design Method"(Dillman 2000; Dillman u. a. 2014) to achieve the highest possible response rate. The effectiveness of this measure was evaluated in two papers (R. Becker und Glauser 2018; R. Becker, Möser u. a. 2019), showing that the incentives contributed to a higher willingness to participate and shorter latency until response after first contact.

For DAB-IV, a total of four surveys are planned. In two main surveys in spring 2020 and 2023 the full mixed-mode design (online questionnaire, CATI, PAPI) is applied. The thematic survey wave 10 on *health and subjective*

well-being was, however, conducted as a single-mode online-only survey in 2021 and an in deph survey on *gender differences and career planning* is planned for 2022 in online-only mode.

Tabelle 1 Overview survey design

	DAB I				**DAB II**	**DAB III**			**DAB IV**			
Wave	E*	1	2	3	4	5	6	7	8	9	10	11
Contact	Classroom Survey				Individual survey							
Survey Mode												
CAWI	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
CATI	✗	✗	✗	✗	✓	✓	✓	✓	✓	✗	✗	✓
PAPI	✓	✗	✗	✗	✗	✗	✓	✓	✗	✗	✗	✓
Incentive	✗	✗	only teachers	✗	10 CHF Migros voucher with ex. split	ball-point pen	10 CHF Migros voucher	10 CHF cash	10 CHF cash	10 CHF cash	10 CHF cash	10 CHF cash

**parent survey, contact via child*

3 Population, sampling and response rate

3.1 Population and sampling[1]

The *target population of the DAB panel study* includes all students from 8th grade (regular classes) of public schools in German-speaking cantons or German-speaking parts of cantons in the 2011-2012 school year [2] The data of the student statistics from the 2009-2010 school year of the Federal Statistical Office's (FSO) were used as the basis for the sampling.

Sampling plan and design: In a first step, the schools of the sample population were assigned to a municipality type according to the typology (9 classes) of the FSO. The use of the municipality typology is reasonable on the one hand due to the correlation of municipality type and social structure. On the other hand, the municipality types reflect the variation of opportunity structures in the education system and on the labour market. In a second step, the number of 8th grades within a school was estimated on the basis of the available data on the number of students at the various schools. Because the data basis for the sampling only contains the total number of students in 8th grade, separated by school type, but not the number of these classes in a school[3]. Classes were sampled on the basis of the cantonal guidelines for minimum and maximum class sizes of the various school types (without selection, school type with basic requirements, school type with extended requirements, pre-gymnasium[4]) using the structural data of the Swiss Conference of Cantonal Ministers of Education (EDK 2011)[5]. Finally, for each of the nine community types, a stratified 10% random sample of the 8th classes was drawn. The stratification took into account the type of school (basic requirements/without selection, extended requirements, pre-gymnasium) and the proportion of foreign-language students within the schools[6]. The number of classes to be drawn within a strata corresponds approximately to the marginal distributions of school types within the community types and the distribution of classes with a low or high proportion of foreign-language students within a school type[7]. After sampling, the school administrators were contacted. If there were several 8th grades of the same school type within a school or if several grades of the same school were drawn during the sampling process, the 8th grade(s) to be interviewed was (were) randomly selected within a school when the schools were contacted[8].

[1] This section was taken from Glauser (2015, S. 125–128).

[2] Before preparing the data for the sampling, the agreement of the cantonal education departments was obtained for the implementation of the research project and for contacting the school administrations. Consent was refused in the cantons of Appenzell Innerrhoden, Schaffhausen, Solothurn and Uri, which is why schools with 8th grade in these cantons are not included in the sample.

[3] Classes with less than 6 students were not included in the sampling.

[4] Pre-gymnasium consists of two year previously to the gymnasium, which provides the Maturitat and access to University.

[5] On the basis of data from the FSO's student statistics, a strict distinction between 8th grades of the school type with extended requirements and the pre-gymnasium is not possible in all cantons. Pre-gymnasium 8th grades are held in the cantons of Aargau, Basel-Land (level P), Basel-Stadt, Freiburg (Orientation School Section A), Glarus, Grisons, Lucerne, Nidwalden, Obwalden, Zug and Zurich. In cantons where pre-gymnasial education is provided by Gymnasiums or cantonal/district schools, allocation is possible without further ado. In the cantons of Basel-Land and Fribourg the number of pre-gymnasium classes within the schools was determined in consultation with the responsible education departments. In the canton of Basel-Land, the data of secondary school students, differentiated by level, from the school year of 2010/11 were used. For the schools in the German-speaking part of the canton of Fribourg, the current figures for the school year of 2011/12 were used. Although the sampling plan only provides for regular classes in public schools, in the Canton of Graubünden private schools that run pre-gymnasial classes were also considered, as adolescents in the Canton of Graubünden usually attend the nearest gymnasium.

[6] In order to consider the proportion of foreign-language students in schools, the median for all schools in the sample was first calculated (15.8%). Schools with a proportion of foreign language speakers below the median were classified as schools with a low proportion, the others as schools with a high proportion of foreign language students.

[7] An exception is the proportion of pre-gymnasium classes in the central municipalities. Although slightly more than 22% of the classes in this type of municipality are classified as pre-gymnasium 8th grade, the proportion of classes to be drawn was limited to 10%. Since with the transition to the 9th grade a non-negligible proportion of students transfer to a gymnasium anyway, this approach prevents the proportion of adolescents from other school types in this type of community from becoming too small.

[8] In the random selection process, first the names of the teachers under consideration were asked and an alphabetical order was formed from the list. A random key was then used to determine which teacher took part in the survey. This procedure could be applied in 84% of the cases of classes to be drawn using the random key. Where school principals did not agree to the random selection, the class(es) proposed

3.2 Sampling and response rates

The sample size, the realised sample as well as the response rates of the first seven survey waves of the DAB panel study are listed in table 2. Overall, the selection procedure described above was used to randomly select a sample of 296 classes at class level from a total of 3045 classes. After contacting school administrators and teachers, 215 out of 296 classes were won for participation in the first survey wave, which corresponds to a response rate of 72.6% at class level. The vast majority of the classes that participated in wave 1 also participated in the following surveys of waves 2 and 3. 12 classes no longer participated in wave 2, while only 4 classes refused to participate in wave 3. The response rate realised at the level of the classes in wave 3 is 67.2% - based on the initial sample of classes in wave 1 (296)[9].

Class lists provided by the teachers were used to prepare the surveys. Based on this information the gross sample of students can be calculated. In this context, a distinction must be made between the DAB sample and a wave-specific gross sample. The DAB sample consists of 4083 students. This sample contains all students that belong to the gross sample in wave 1 (3815) as well as those students that entered a class of the DAB sample at a later point in time (205 before wave 2 and 63 before wave 3). The wave-specific gross samples refer to those students that are eligible to participate in wave X, i.e. these students are part of the DAB sample and have not yet been definitively eliminated from the sample at the time of wave X. Thus, the gross sample (N=3997) in wave 2 is composed of the gross sample in wave 1 (N=3815), plus the students (+205) newly entered in classes of the DAB sample from wave 2 onwards, minus the students (-23) definitely eliminated in wave 1.

Generally the wave-specific gross sample corresponds to the number of eligible students in wave X. Whereby wave 4 is an exception, as 245 people who for various reasons had not participated in wave 3 were inadvertently not invited to take part in the survey.

With the switch from classroom (waves 1 - 3) to individual surveys from wave 4 onwards, *contactability* (by post, electronically or by telephone) of the individuals in the DAB sample is a prerequisite for participation in the survey. Thus, the gross sample in wave 4 includes N= 3526, but only 3281 can be described as eligible to participate due to the explanations given above. Of these 3281 individuals, contact information of 2655 could be used in wave 4. From wave 4 onwards, various measures were taken to complete missing contact details of respondents. As a result, the number of individuals who could be contacted slightly increased or was kept relatively constant in the following waves (wave 5: 2800, wave 6: 2720, wave 7: 2489).

Closely related to the contactability of respondents is the number of people who *definitively dropped out*. While only 23 people definitively dropped out of the DAB sample after the first wave, this number increased to a total of 348 in wave 2 (+325) and 557 in wave 3 (+209). The strongest increase in definitive dropouts occurred in wave 4 (+662) to a total of 1219 people. This is due to a lack of contact details on the one hand and on the other hand to a lower willingness to participate in the individual survey. The number of definite dropouts is low in waves 5 (+126), 6 (+242), and 7 (+96), and amounts to a total of 1683 people by wave 8.

by the school principals were included in the sample. Due to the very low proportion of classes that could not be randomly selected, no systematic bias of the sample can be assumed.

[9] After finishing the 8th grade, about 700 students have changed their class and/or school. From the second survey onwards, these individuals were contacted personally, if possible, and invited to participate in the survey. During the first survey, the respondents were asked for their contact information. In addition, towards the end of the 8th grade, all teachers were sent address forms for students who left the class and/or school. In both cases it was not compulsory to fill in or return the contact details. A large part of the sample drop-out at class level as well as among the individuals in waves 2 and 3 is due to changes of class or school. The proportion of refusals by teachers (3 classes) is significantly lower than that resulting from the merging or closing of classes (9 classes).

Tabelle 2 DAB: Overview of the sample and response rates

	Wave 1 Jan./Feb. 2012	Wave 2 Aug./Oct. 2012	Wave 3 May/Jun. 2013	Wave 4 Oct./Nov. 2014	Wave 5 Jun./Aug. 2016	Wave 6 May/Jun. 2017	Wave 7 May/Jun. 2018	Wave 8 May/Jun. 2020
Sample size								
Gross sample (classes)	296	215	203	-	-	-	-	-
Net sample (classes)	215	203	199	-	-	-	-	-
DAB sample	4083	4083	4083	4083	4083	4083	4083	4083
Gross sample (students)	3'815	3'997	3'735	3'526	2'864	2'738	2'496	2490
Eligible to participate	3'815	3'997	3'735	3'281	2'864	2'738	2'496	2490
Not yet in class in wave X	205	63	-	-	-	-	-	-
Contactable students	-	-	-	2'655	2'800	2'720	2'489	2464
Def. drop out in wave X	23	325	209	662	126	242	96	n.a.
Total drop out until walle X	23	348	557	1219	1345	1587	1683	n.a.
Realised answers								
Students	3'680	3'331	3'281	2'236	2'229	2'061	1'958	2009
Online (*in class*)	3'680	3110	3089	-	-	-	-	-
Online (*individual*)	-	221	192	1'227	1'330	1'375	1'646	1877
CATI	-	-	-	1'009	899	598	287	132
PAPI	-	-	-	-	-	88	25	-
Parents	2'284	-	-	-	-	-	-	-
Online	404	-	-	-	-	-	-	-
PAPI	1'880	-	-	-	-	-	-	-
Response rates in %								
Classes in wave (gross / realized)	72.6%	94.4%	98.0%	-	-	-	-	-
Total classes	72.6%	68.6%	67.2%	-	-	-	-	-
students in wave X (gross / realised)	96.5%	83.3%	87.8%	63.4%	77.8%	75.3%	78.4%	80.7%
Contactable students in wave X	-	-	-	84.2%	79.6%	75.8%	78.7%	82.4%
students total (DAB sample / realized)	94.9%	82.9%	80.4%	54.8%	54.6%	50.5%	48.0%	49.2%
Parents (gross / realised)	59.9%	-	-	-	-	-	-	-

Realised responses: In wave 1, 3680 individuals took part in the survey. In relation to the gross sample, a response rate of 96.5% was achieved. In the two subsequent surveys, the response rate - with reference to the wave-specific gross sample - is above 80% (wave 2: 83.3%, wave 3: 87.8%). As mentioned above, the strongest decrease in the number of realised responses occurs in wave 4. Although 2236 individuals took part in this survey, in relation to the gross sample (= 3281) only a response rate of 63.4% was achieved. In waves 5 to 8, again higher response rates were achieved (wave 5: 77.8%, wave 6: 75.3%, wave 7: 78.4%, wave 8: 80.7%). If the response rates from wave 4 are calculated in relation to the sample of contactable individuals, the following response rates were achieved: (wave 4: 84.2%, wave 5: 79.6%, wave 6: 75.8%, wave 7: 78.7%, wave 8: 82.4%).

With regard to the *response rates regarding the DAB sample* (N = 4083), the following picture emerges. In wave 1 a response rate of 94.9% was achieved[10]. Due to class and/or school changes after completion of the 8th grade, the response rate decreases to 82.9% (wave 2) and 80.4% (wave 3). The strongest decline is seen in wave 4 and thus in the first individual survey after grade 9 and the end of compulsory schooling. In wave 4, for example, a response rate of 54.8% was still achieved. Fortunately, the response rate after wave 4 was kept relatively constant (wave 5: 54.6%, wave 6: 50.5%, wave 7: 48.0%, wave 8: 49.2%)).

For economic reasons, the *parents* were interviewed only in wave 1. With reference to the gross sample in wave 1 (N = 3815) 2284 answers could be realized. This corresponds to a response rate of 59.9%. 82% of the parents returned the PAPI questionnaire, while 18% completed the questionnaire online.

[10]To calculate the response rate with reference to the DAB sample, a total of 268 people are subtracted in wave 1 who had not yet entered a class of the DAB sample in wave 1 (W2: 205 students, W3: 63 students) and therefore could not participate in the survey. A similar approach was used to calculate the response rate in wave 2.

4 Longitudinal weights – Construction and implementation

The need to weight the data comes from the inherent problem of dropouts and non-response across the subsequent measurement points in panel studies. The inferential statistical weights take the dropout or refusal into account from one survey wave to the next. As is usual in panel studies, the weights are constructed as the reciprocal of the product of the participation probabilities of the individual survey waves. For the DAB panel study this results in:

$$G_{t_i} = \frac{N_{\text{Stratum}_i}}{n_{\text{Stratum}_i}} \times \frac{1}{P_{t1_i}} \times \ldots \times \frac{1}{P_{tX_i}} \tag{1}$$

with:

G_{t_i} = Panel weight at time of survey t for respondent i

$\frac{N_{\text{Stratum}_i}}{n_{\text{Stratum}_i}}$ = Total classes within stratum of i / number of classes realised within stratum of i /

$\frac{1}{P_{t1_i}}$ = Probability of participation in DAB survey wave **t1** for respondent i

$\frac{1}{P_{tX_i}}$ = Probability of participation in DAB survey wave **tX** for respondent i

The provided weights take the sample design into account: the ratio of the total number of classes within the formed strata of observation i and the drawn and realised classes within each stratum of observation i are included in the calculation of the weights. Additionally, the participation probabilities of the individual surveys are included in the weights as reciprocal values of the product of the wave-specific participation probabilities. For example, the panel weight at survey time $t3$ for respondent i is made up of:

$$G_{t_i} = \frac{N_{\text{Stratum}_i}}{n_{\text{Stratum}_i}} \times \frac{1}{P_{t1_i}} \times \frac{1}{P_{t2_i}} \times \frac{1}{P_{t3_i}}$$

The weights of the other survey waves are generated in the same way. In the following, it is explained how the wave-specific probabilities of participation were calculated (Chapter 4.1). In addition, chapter 4.2 explains the truncation of the wave-specific gross weights using the weights from wave 5. Finally, chapter 4.3 contains an overview of the weights provided in the DAB data and chapter 4.4 describes the use of sample weights in statistical analyses.

4.1 Probability of participation

4.1.1 Probability of participation in wave 1 and 2

Since no information from previous surveys or external data sources are available to estimate the *probability of participation in wave 1*, the probability of participation is based on the average participation of individuals within their stratum. The average probability of participation is .901, with the range of values between min = .742 and max = .978 (see Table 3)).

The same procedure is applied to the calculation of the *probability of participation in wave 2*, since no information from the first survey is available for the estimation of the probability of participation in wave 2 for those who entered in a class of the DAB sample from wave 2 onwards. The sample size is N = 4060 minus the number of people who definitively dropped out from wave 1 (= 23), and the average probability of participa-tion in wave 2 is .820, with the range of values between min. = .5 and max. = 1 (see Table 3).

Tabelle 3 Probability of participation in wave 1 and 2

	mean	sd	min	max	N
Probability of participation W1	.9012981	.0508371	.7419355	.9782609	4083
Probability of participation W2	.8204434	.0694607	.5	1.0	4060

4.1.2 Probability of participation in wave 3

From wave 3 onwards, the estimation of the probability of participation is based on the results of logistic regressions. Subtracting individuals who definitively dropped out up to and including wave 2 (= 348), the sample size in wave 3 is N = 3735. The following variables are used in the logistic regression model to the dependent variable, which has the values 0 (no participation) or 1 (participation). The municipality type of the classes in the sample is included using a categorical variable, whereby municipality types are partially summarized: central municipality/suburban municipality (= reference category), high-income periurban municipality, touristic municipality, industrial-tertiary municipality, rural commuter/agricultural mixed/agricultural municipality. Additionally, it is controlled for gender (with men as reference category), type of school (basic requirements/without selection (reference category), extended requirements, pre-gymnasium) and the participation mode in wave 2 (classroom survey (reference category), individual survey, without participation in W2).

The results of the logistic regression on the *probability of participation in wave 3* indicate that individuals from the community types rural-commuter/agricultural-mixed/agricultural have a higher chance of participating in wave 3 than those from the reference category (central community/suburban community). There are no differences in participation by gender in wave 3. In contrast, adolescents from pre-gymnasium have a higher chance of participating in wave 3 than adolescents from school types with basic requirements, while the latter do not differ significantly from adolescents who attended school types with extended requirements. Finally, the chance that an individual participates in wave 3, given that they either participated in the individual survey in wave 2 or did not participate at all, is significantly lower compared to those who participated in the classroom survey in wave 2.

Tabelle 4 Participation in wave 3, logistic regression, odds ratios

	OR	SE	z	p
Municipality type (*Ref.*: Center Community/Suburban)	1.000	.	.	.
High income / Periurban	0.900	0.408	−0.233	0.816
Touristic	0.945	0.327	−0.164	0.869
Industrial tertiary	1.342	0.270	1.457	0.145
Rural commuters/agricultural mixed/agricultural	2.038**	0.494	2.937	0.003
Women (*Ref.*: Men)	0.972	0.107	−0.258	0.797
School type 8th grade (*Ref.*: Basic requirements)	1.000	.	.	.
Advanced requirements	1.322	0.260	1.416	0.157
Pre-baccalaureate	1.832*	0.439	2.524	0.012
Participation mode in t2 (*Ref.*: in class)	1.000	.	.	.
Individual survey	0.150***	0.031	−9.110	0.000
Without participation	0.209***	0.032	−10.076	0.000
Beobachtungen	3735			
Pseudo R^2	0.095			
log-likelihood	−1251.0			
$\chi^2(df)$	222.1(9)			

The probability of participation for observation i is estimated using the effect coefficients and the individual variable values. Descriptive values of the probability of participation in wave 3 are shown in Table 5.

Tabelle 5 Probability of participation in wave 3

	mean	sd	min	max	N
Probability of participation W3	.8784471	.0965799	.526704	.9694722	3735

4.1.3 Probability of participation in wave 4

In wave 4, separate models are estimated for contactability and for participation. As explained above, the strongest decrease of the response rate was observed when switching from classroom surveys to individual surveys in wave 4. Therefore, the probability of contactability as well as the probability of participation was considered for the calculation of the longitudinal weight for wave 4.

In addition to the control variables reported for wave 3, the status of educational decision in wave 3 (occupational training (reference category), secondary school, Interim solution/bridge-year courses, etc., no information), the social class situation of the parents (EGP classes I/II (reference category), EGP classes III/IV, EGP classes V-VII, EGP class missing), the time preference in wave 3 (z-standardised) as well as the response time in wave 3 are also taken into account to estimate the *contactability in wave 4* (see table 6). A positive aspect is that the contactability in wave 4 is neither related to gender, nor to the type of attended school or the social background. This also applies to a limited extent to the status of the educational decision in wave 3. Only respondents, who had stopped the questionnaire before the questions on the status of their training decision, had a lower chance of being contacted in wave 4 than those seeking occupational training. Nonetheless, it appears that those with a strong time preference for the present, i.e. young people for whom it is important that their training is as short as possible, have a lower chance of being contacted in wave 4. The same applies to those who did not participate in wave 2 compared to people who had participated in the classroom survey. Finally, it turns out that young people

who took more time to answer the questionnaire in wave 3 have a higher chance of being contacted in wave 4. Descriptive parameters of the probability of being contacted in wave 4 are shown below.

Tabelle 6 Contactability in wave 4, logistic regression, odds ratios

	OR	SE	z	p
Municipality type (*Ref.*: Center Community/ Suburban)	1.000	.	.	.
High income/Periurban	1.488	0.441	1.341	0.180
Touristic	2.887*	1.276	2.398	0.016
Industrial tertiary	1.128	0.287	0.476	0.634
Rural commuters/agricultural mixed/ agricultural	1.494	0.462	1.301	0.193
Women (*Ref.*: Men)	1.196	0.130	1.648	0.099
School type 8th grade (*Ref.*: Basic requirements)	1.000	.	.	.
Advanced requirements	0.893	0.189	−0.537	0.591
Pre-baccalaureate	1.488	0.538	1.100	0.272
Status educational decision in t3 (*Ref.*: VET)	1.000	.	.	.
(Middle) school	1.047	0.187	0.259	0.796
Interim solution/bridge-year courses, etc.	0.800	0.123	−1.449	0.147
No information	0.214***	0.058	−5.727	0.000
Social class of parents (*Ref.*: EGP-Klasse I/II)	1.000	.	.	.
III/IV	0.970	0.167	−0.178	0.859
V-VII	0.824	0.116	−1.372	0.170
No information	0.739	0.133	−1.674	0.094
Time preferences in t3, standardised (t3timepref1)	0.861**	0.039	−3.272	0.001
Participation mode in t2 (*Ref.*: in class)	1.000	.	.	.
Individual survey	1.452	0.356	1.521	0.128
Without participation	0.482***	0.073	−4.841	0.000
Response time in t3 (t3duration)	1.000*	0.000	1.983	0.047
Observation	3281			
Pseudo R^2	0.056			
log-likelihood	−1510.2			
$\chi^2(df)$	126.6(17)			

Tabelle 7 Probability of contactability wave 4

	mean	sd	min	max	N
Contactability W4	.8092045	.0982937	.1168456	.969012	3281

With regard to the *probability of participation in wave 4* the following is found (see Table 8): In wave 4, women have a higher chance of participating in the survey than men. Individuals who attended a type of school with extended requirements or a pre-baccalaureate instead of the type of school with basic requirements in grade 8 also have a higher chance of participation. Furthermore, individuals with very good mathematics scores (grade 5.5 or 6) have a higher chance of participating than those with average grades in mathematics (grades 4 to 5), too. In terms of class level, adolescents from working class families (EGP classes V-VII) do not have a lower chance of participation than those from upper and lower service classes (EGP classes I/II). No effect on participation is observed in relation to the highest level of education of parents.

There are no significant effects regarding the parents' country of birth or the generational status of the child. Finally, adolescents who consider it important for a man to earn a lot have a lower chance of participation. This is not a gender effect. Descriptive parameters of the probability of participation in wave 4 are shown below.

Tabelle 8 Participation in wave 4, logistic regression, odds ratios

	OR	SE	z	p
Municipality type (*Ref.*: Center Community/ Suburban)	1.000	.	.	.
High income/Periurban	1.313	0.286	1.246	0.213
Touristic	0.515***	0.098	−3.476	0.001
Industrial tertiary	1.022	0.191	0.117	0.907
Rural commuters/agricultural mixed/ agricultural	0.833	0.146	−1.041	0.298
Women (*Ref.*: Men)	1.344**	0.143	2.772	0.006
School type 8th grade (*Ref.*: Basic requirements)	1.000	.	.	.
Advanced requirements	1.919***	0.270	4.636	0.000
Pre-baccalaureate	2.149**	0.539	3.050	0.002
Mathematics scores 9th class (*Ref.*: average)	1.000	.	.	.
No information	0.642	0.250	−1.139	0.255
Below average	0.822	0.122	−1.322	0.186
Above average	1.732**	0.321	2.960	0.003
Status educational decision in t3 (*Ref.*: VET)	1.000	.	.	.
(Middle) school	1.439	0.283	1.855	0.064
Interim solution/bridge-year courses, etc.	1.151	0.215	0.750	0.453
Social class of parents (*Ref.*: EGP class I/II)	1.000	.	.	.
III/IV	1.431*	0.257	1.999	0.046
V-VII	1.180	0.211	0.927	0.354
No information	1.009	0.237	0.037	0.970
Highest level of education parents (*Ref.*: Max. ISCED 3C) 1.000	1.000	.	.	.
ISCED 3B	1.245	0.244	1.121	0.262
ISCED 3A	1.467	0.394	1.426	0.154
ISCED 4-6	1.420	0.300	1.659	0.097
No information	1.003	0.237	0.011	0.991
Country of birth parents (*Ref.*: CH)	1.000	.	.	.
EU/EFTA - CH-mixed rest	1.271	0.634	0.481	0.630
Balkan, TUR, POR	1.047	0.529	0.090	0.928
No information	1.022	0.387	0.057	0.954
Generational status (*Ref.*: 3rd generation)	1.000	.	.	.
1st generation	0.492	0.267	−1.305	0.192
2nd generation	0.709	0.368	−0.663	0.507
2.5 Generation	0.558	0.275	−1.185	0.236
No information	0.514	0.236	−1.448	0.148
Importance of status maintenance (t3impjstat3)	1.090	0.059	1.593	0.111
Important, that men earn a lot (t3malrole2)	0.857*	0.058	−2.269	0.023
Observations	2655			
Pseudo R^2	0.069			
log-likelihood	−1077.3			
$\chi^2(df)$	183.8(28)			

Tabelle 9 Probability of participation in wave 4

	mean	sd	min	max	N
Probability of participation W4	.8421846	.0929037	.3319636	.9766418	2655

4.1.4 Probability of participation in wave 5

The calculation of the *probability of participation in wave 5* is based on the logistic regression model shown in Table 10. Wave 5 also shows a gender effect. Women show a higher willingness to participate than men. Likewise, the chance of participation in wave 5 is lower for adolescents who had attended a type of school with basic requirements instead of a type of school with extended requirements or a pre-school. With regard to social background (occupational situation and education of parents), no direct effects can be seen. These are mostly

mediated by the type of school attended in grade 8. Also without direct effects is the country of birth of the parents and the generation status of the children.

Finally, with regard to the training situation and participation in wave 4, the following is found: Those for whom no information on the training situation is available from wave 4 and those who participated in the telephone survey in wave 4 have a lower chance of participating in wave 5. Descriptive parameters for the probability of participation in wave 5 are shown in Table 11.

Tabelle 10 Participation in wave 5, logistic regression, odds ratios

	OR	SE	z	p
Municipality type (*Ref.*: Center Community/ Suburban)	1.000	.	.	.
High income/Periurban	0.973	0.169	−0.155	0.876
Touristic	1.234	0.410	0.633	0.527
Industrial tertiary	1.134	0.171	0.831	0.406
Rural commuters/agricultural mixed/ agricultural	1.311	0.216	1.646	0.100
Women (*Ref.*: Men)	1.485***	0.155	3.784	0.000
School type 8th grade (*Ref.*: Basic requirements)	1.000	.	.	.
Advanced requirements	1.770***	0.204	4.962	0.000
Pre-baccalaureate	2.451***	0.630	3.486	0.000
German score 9th class (*Ref.*: average)	1.000	.	.	.
No information	1.122	0.263	0.492	0.623
Below average	0.731	0.144	−1.595	0.111
Above average	1.268	0.210	1.431	0.153
Status educational decision in t3 (*Ref.*: VET)	1.000	.	.	.
Secondray school	1.018	0.211	0.087	0.930
Interim solution/bridge-year courses, etc.	1.019	0.165	0.118	0.906
No information	2.554**	0.759	3.156	0.002
Social class of parents (*Ref.*: EGP class I/II)	1.000	.	.	.
III/IV	0.867	0.144	−0.862	0.389
V-VII	1.027	0.162	0.168	0.867
No information	0.836	0.168	−0.889	0.374
Highest level of education parents (*Ref.*: Max. ISCED 3C)	1.000	.	.	.
ISCED 3B	1.275	0.195	1.587	0.113
ISCED 3A	1.245	0.268	1.018	0.309
ISCED 4-6	1.234	0.207	1.255	0.209
No information	1.787**	0.387	2.677	0.007
Country of birth parents (*Ref.*: CH)	1.000	.	.	.
EU/EFTA - CH-mixed rest	0.651	0.291	−0.959	0.338
Balkan, TUR, POR	0.601	0.283	−1.081	0.280
No information	1.405	0.471	1.015	0.310
Generational status (*Ref.*: 3rd generation)	1.000	.	.	.
1st generation	0.765	0.381	−0.538	0.591
2nd generation	1.107	0.498	0.226	0.821
2.5 generation	1.110	0.510	0.228	0.820
No information	0.863	0.241	−0.374	0.709
Educational situation in t4 (*Ref.*: VET)	1.000	.	.	.
(Middle) school	1.275	0.319	0.969	0.332
Interim solution/bridge-year courses, etc.	0.912	0.301	−0.279	0.780
No information	0.286**	0.118	−3.036	0.002
Participation mode in t4 (*Ref.*: online)	1.000	.	.	.
CATI	0.638***	0.079	−3.620	0.000
No participation	0.531	0.228	−1.477	0.140
Observations	2864			
Pseudo R^2	0.135			
log-likelihood	−1311.1			
$\chi^2(df)$	394.9(32)			

Tabelle 11 Probability of participation in wave 5

	mean	sd	min	max	N
Probability of participation W5	.7782821	.1633423	.1415374	.9755289	2864

4.1.5 Probability of participation in wave 6

In contrast to the previous waves 4 and 5, there is no gender difference with regard to the chance of participating in wave 6 (see table 12 on page 18). The effects for the type of school attended at lower secondary level are unchanged, i.e. young people from the school type with basic requirements are less willing to participate compared to young people from the other school types (extended requirements, pre-baccalaureate). On the other hand, under control of the type of school, no direct effect of the grades in German and mathematics can be observed. Rather unexpected is the finding that adolescents from parents belonging to the working classes (EGP classes V-VII) have a higher chance of participating in wave 6 than adolescents in the reference category (EGP classes I/II: upper and lower service classes). Moreover, it is not shown that children parents who have a tertiary education have a higher chance of participating in wave 6. This is also true when tertiary educational qualifications are used as a reference category. While there are no significant effects with regard to the country of birth of the parents, individuals of the 2nd generation participate in wave 6 with a lower chance than those of the reference category (3rd generation). Finally, as in wave 5, it is apparent that those who participated in the previous wave of the telephone survey have a lower chance of participating in wave 6. Descriptive parameters of the probability of participation in wave 6 are shown in Table 13 (see page 18).

4.1.6 Probability of participation in wave 7

As in wave 6, there is no gender difference in wave 7 with regard to the chance of participating in wave 7 (see table 14 on page 14). The effects for the type of school attended at lower secondary level remain un-changed. While there are no significant effects of the German grade, an insufficient grade reduces the chance of participation in wave 7. Moreover, the chance of participation is increased for individuals who wanted to start secondary school in wave 3, while no direct effects on the classroom situation and the educational level of the parents can be observed. Furthermore, it can be seen that respondends who started post-secondary education in wave 6 have a higher chance of participating in wave 7 than people in basic occupational training. Finally, the chance of participation in wave 7 is related to the political interest - observed in wave 6. The more pronounced the political interest, the higher the chance that people took part in the seventh survey of the DAB panel study. Descriptive parameters of the probability of participation in wave 7 are shown in Table 15 (see page 19).

4.1.7 Probability of participation in wave 8

There is no gender difference regarding the probability to participate in wave 8 (see table 16 on page 16). The effects of the type of school attended at lower secondary level are strong; those who attended a pre-baccalaureate school have a significantly higher chance of participating in wave 8. While there are no significant effects of the German grade, a very good mathematics grade increases the chance of participation in wave 8. While no direct effects can be observed for migration background and class situation, the chance of participation increases with a

higher educational level of the parents. Furthermore, it can be seen that individuals who started post-secondary education in wave 7 have a higher chance of participating in wave 8 than those who are still in basic vocational

Tabelle 12 Participation in wave 6, logistic regression, odds ratios

	OR	SE	z	p
Women (*Ref.*: Men)	1.117	0.124	0.995	0.32
School type 8th grade (*Ref.*: Basic requirements)	1.000	.	.	.
Advanced requirements	1.536***	0.193	3.413	0.001
Pre-baccalaureate	2.074**	0.520	2.907	0.004
German scores 9th class (*Ref.*: average)	1.000	.	.	.
No information	1.560	1.641	0.423	0.672
Below average	0.915	0.235	−0.347	0.729
Above average	1.178	0.239	0.809	0.418
Mathematics scores 9th class (*Ref.*: average)	1.000	.	.	.
No information	0.593	0.611	−0.507	0.612
Below average	0.962	0.159	−0.237	0.813
Above average	1.200	0.193	1.136	0.256
Status educational decision in t3 (*Ref.*: VET)	1.000	.	.	.
(Middle) school	1.171	0.263	0.702	0.483
Interim solution/bridge-year courses, etc.	1.078	0.197	0.411	0.681
No information	0.820	0.202	−0.805	0.421
Social class of parents (*Ref.*: EGP class I/II)	1.000	.	.	.
III/IV	1.189	0.229	0.899	0.369
V-VII	1.427*	0.245	2.072	0.038
No information	1.100	0.269	0.389	0.697
Highest level of education parents (*Ref.*: Max. ISCED 3C)	1.000	.	.	.
ISCED 3B	0.625*	0.118	−2.479	0.013
ISCED 3A	0.616*	0.137	−2.174	0.030
ISCED 4-6	0.795	0.161	−1.131	0.258
No information	0.751	0.176	−1.221	0.222
Country of birth parents (*Ref.*: CH)	1.000	.	.	.
EU/EFTA - CH-mixed rest	1.275	0.590	0.524	0.600
Balkan, TUR, POR	1.052	0.481	0.110	0.912
No information	0.932	0.335	−0.197	0.844
Generational status (*Ref.*: 3rd generation)	1.000	.	.	.
1st generation	0.630	0.292	−0.997	0.319
2nd generation	0.421*	0.186	−1.961	0.050
2.5 generation	0.642	0.305	−0.932	0.351
No information	0.371*	0.150	−2.455	0.014
Status educational decision in t5 (*Ref.*: VET)	1.000	.	.	.
(Middle) school	1.170	0.290	0.633	0.527
No information	0.538*	0.144	−2.316	0.021
General satisfaction in t5 (*Ref.*: rather unsatisfied)	1.000	.	.	.
Rather satisfied	1.437	0.318	1.638	0.101
No information	1.527	1.956	0.331	0.741
Participation mode in t5 (*Ref.*: online)	1.000	.	.	.
CATI	0.446***	0.060	−6.001	0.000
No participation	0.101	0.127	−1.813	0.070
Observations	2738			
Pseudo R^2	0.240			
log-likelihood	−1163.9			
χ^2 (*df*)	678.3(36)			

Note: controlled, but not shown: odds ratios of municipality type.

Tabelle 13 Probability of participation in wave 6

	mean	sd	min	max	N
Probability of participation W6	.7527392	.2249789	.0836424	.9745457	2738

Tabelle 14 Participation in wave 7, logistic regression, odds ratios

	OR	SE	z	p
Women (*Ref.*: Men)	1.228	0.156	1.613	0.107
School type 8th grade (*Ref.*: Basic requirements)	1.000	.	.	.
Advanced requirements	1.649***	0.207	3.980	0.000
Pre-baccalaureate	2.380***	0.601	3.431	0.001
German scores 9th class (*Ref.*: average)	1.000	.	.	.
No information	0.458	0.264	−1.354	0.176
Below average	1.286	0.329	0.984	0.325
Above average	1.101	0.241	0.442	0.659
Mathematics scores 9th class (*Ref.*: average)	1.000	.	.	.
No information	1.760	0.985	1.010	0.313
Below average	0.693*	0.122	−2.088	0.037
Above average	0.927	0.183	−0.383	0.702
Status educational decision in t3 (*Ref.*: VET)	1.000	.	.	.
(Middle) school	1.587*	0.350	2.093	0.036
Interim solution/bridge-year courses, etc.	0.987	0.217	−0.061	0.951
No information	1.009	0.278	0.031	0.975
Social class of parents (*Ref.*: EGP class I/II)	1.000	.	.	.
III/IV	1.119	0.249	0.504	0.614
V-VII	0.770	0.159	−1.266	0.206
No information	0.918	0.252	−0.31	0.756
Highest level of education parents (*Ref.*: Max. ISCED 3C)	1.000	.	.	.
ISCED 3B	1.188	0.213	0.963	0.335
ISCED 3A	1.356	0.367	1.125	0.261
ISCED 4-6	1.352	0.291	1.403	0.161
No information	0.943	0.240	−0.23	0.818
Country of birth parents (*Ref.*: CH)	1.000	.	.	.
EU/EFTA - CH-mixed rest	0.744*	0.110	−1.994	0.046
Balkan, TUR, POR	0.818	0.156	−1.054	0.292
No information	0.782	0.185	−1.038	0.299
Status educational decision in t6 (*Ref.*: VET)	1.000	.	.	.
(Middle) school	1.442	0.436	1.208	0.227
Paid Work	0.861	0.154	−0.839	0.402
Post secondary training	3.116**	1.315	2.693	0.007
Other	1.062	0.242	0.262	0.793
No participation	0.460**	0.127	−2.802	0.005
General interest in politics in t6 (*Ref.*: low interest)	1.000	.	.	.
Partly	1.992***	0.350	3.917	0
High interest	2.560***	0.557	4.326	0
No participation	0.299***	0.075	−4.836	0
Observations	2496			
Pseudo R^2	0.255			
log-likelihood	−968.8			
χ^2 (*df*)	646.4(34)			

Note: controlled, but not shown: odds ratios of municipality type.

Tabelle 15 Probability of participation in wave 7

	mean	sd	min	max	N
Probability of participation W7	.7844551	.2196565	.1377769	.9940798	2496

education. However, the previously made training decision in t3 shows no effect. Finally, it can be seen that people who are satisfied with their educational and occupational history are more likely to participate in the eighth survey of the DAB panel study than those who are rather dissatisfied. Descriptive characteristics of the probability of participation in wave 8 are shown in table 17 (see page 20).

Tabelle 16 Participation in wave 8, logistic regression, odds ratios

	OR	SE	z	p
Women (*Ref.*: Men)	1.304	0.184	1.878	0.060
School type 8th grade (*Ref.*: Basic requirements)	1.000	.	.	.
Advanced requirements	1.925***	0.298	4.234	0.000
Pre-baccalaureate	5.275***	2.073	4.230	0.000
German grade 9th class (*Ref.*: average)	1.000	.	.	.
No information	3.356	2.652	1.532	0.125
Below average	1.308	0.416	0.844	0.399
Above average	1.068	0.245	0.285	0.776
Mathematics grade 9th class (*Ref.*: average)	1.000	.	.	.
No information	0.363	0.276	−1.331	0.183
Below averag	0.867	0.178	−0.695	0.487
Above average	2.263***	0.499	3.699	0.000
Status educational decision in t3 (Ref.: VET)	1.000	.	.	.
(Middle) school	0.919	0.216	−0.360	0.719
Interim solution/bridge-year courses, etc.	1.120	0.249	0.511	0.610
No information	1.318	0.454	0.803	0.422
Social class of parents (*Ref.*: EGP-Klasse I/II)	1.000	.	.	.
III/IV	1.102	0.254	0.419	0.675
V-VII	1.263	0.273	1.079	0.281
No information	1.492	0.421	1.420	0.156
Highest level of education parents (*Ref.*: Max. ISCED 3C)	1.000	.	.	.
ISCED 3B	1.697*	0.359	2.501	0.012
ISCED 3A	1.659*	0.425	1.977	0.048
ISCED 4-6	1.727*	0.432	2.182	0.029
No information	1.119	0.303	0.416	0.678
Country of birth parents (Ref.: CH)	1.000	.	.	.
EU/EFTA - CH-mixed other	0.839	0.153	−0.962	0.336
Balkan, TUR, POR	0.754	0.157	−1.353	0.176
No information	0.599	0.158	−1.940	0.052
Status educational situation in t7 (Ref.: VET)	1.000	.	.	.
(Middle) school	2.569	1.657	1.464	0.143
Paid Work	1.134	0.299	0.478	0.633
Post secondary education	2.290*	0.859	2.208	0.027
Other	1.044	0.307	0.145	0.885
No information	0.403	0.249	−1.473	0.141
Satisfaction with educational trajectory (Ref.: unsatifsfied)	1.000	.	.	.
Satifsfied	1.726*	0.421	2.240	0.025
No information	0.319	0.195	−1.871	0.061
Observations	2400			
Pseudo R^2	0.295			
log-likelihood	−751.6			
χ^2 (*df*)	536.8			

Note: controlled, but not shown: odds ratios of municipality type

Tabelle 17 Probability of participation in wave 8

	mean	sd	min	max	N
Probability of participation W8	.8370833	.2060078	.0821911	.9973385	2400

4.2 Truncation of the raw weights [11]

With the estimation of the participation probabilities for all survey waves, all wave-specific probabilities for the calculation of the weighting variables according to equation 1 are available. Given that the systematic default

[11] This section has been adopted in consultation and with the consent of Stefan Sacchi (2011) and adapted to the DAB data.

processes underlying the non-response can be approximated by the estimation models used, the provided weighting variables allow for unbiased sample estimates.

In the application of sample weights, there is a conflict between the aim of correcting non-response bias and minimizing the negative impact of the weighing on the precision of sample-based estimates and projections. Basically, the expected decrease in estimation precision increases with the variance of the weighting variables. In panel data, the dispersion of weights increases from wave to wave due to the relationship given in equation 1. This affects the estimation precision over the survey waves or the estimation precision decreases. However, it is often the case that only a few observations have very high weights and that these weights strongly influence the dispersion of the weighting variables, which decreases the estimation precision. Few extreme weights are also problematic because they not only negatively influence the estimation precision, but also because this applies in particular to evaluations that relate to smaller subsamples. According to Kish (1992), for example, the variance of a weighted mean estimation (μ_W) iincreases in comparison with an unweighted mean estimation (μ) according to the following expression, where cv corresponds to the coefficient of variation of the weighting variable used:

$$\text{var}(\mu_W) = \text{var}(\mu) * (1 + \text{cv}^2) \tag{2}$$

The disadvantages of extreme weights mentioned above can be avoided or at least mitigated by truncating the weights. All individual weights that exceed a defined upper limit are truncated to this limit. The optimal benchmark for the truncation is determined by an evaluation based on the relationship in equation 2. As an example, this is shown in Table 15 for the truncation of the weights of wave 5, where the used weighting variable from wave 5 - according to equation 1 - was recalibrated to an average value of 1. For this purpose the gross weight of wave 5 was divided by its mean value [12]. The recalibration has no effect on the optimisation of the truncation. The first column shows the systematically varied upper benchmark of the calibrated weights, on which the weights from wave 5 are truncated. The second column shows the resulting coefficient of variation of the individual weights of wave 5 that are truncated to different parameters. The third column provides information on how the variance of the weighted sample estimator according to equation 2 changes depending on the selected benchmark of the truncation. Without the truncation, a purely weighting-based reduction of the estimation accuracy by a factor of about 1.4 would be expected. The more the variability of the weighting variables is reduced by means of truncation, the more the estimation precision decreases.

If the number of individual weights affected by the support is also included (fourth column), support at a benchmark of 3.5 proves to be optimal overall in the selected example: With a more radical support, the precision of the sample estimation increases. However, this would result in a strong increase in the number of individual weights affected by the support, which would improve the efficiency of the non-response corrections and the predictability of the values would diminish. The recalibrated gross weights of wave 5 considered in the example are truncated at an upper corner value of 3.5.

The estimation accuracy can be improved thanks to the truncation of weights. Table 19 shows how the estimation precision of the individual waves is improved by the truncation and how many extreme weights are

[12] It should be noted that the calibration to the mean value of 1 is only performed with the sample of the people who participated in the corresponding wave. Therefore, the case numbers of the weighting variables provided in the DAB data, as shown in chapter 4.3, differ. While the gross weight is available for all people who are considered eligible to participate in wave X (see Table 1), the calibrated and truncated weights are only available for observations that participated in wave X.

Tabelle 18 Truncation of the calibrated gross weights from wave 5

N	2229		
truncated Gewichte	42	in %	1.89%
	cv	var(μ_W)/var(μ)	number
Without truncation	0.6569	1.4315	0
Truncation of the recalibrated weights from …			
>8	0.6555	1.4297	1
>7	0.6513	1.4242	1
>6	0.6439	1.4146	3
>5	0.6324	1.3999	8
>4.5	0.6190	1.3832	14
>4	0.6033	1.3640	17
>3.5	0.5735	1.3289	42
>3	0.5341	1.2853	54
>2.5	0.4901	1.2402	126

affected by the truncation. [13] Generally, it must be noted that the truncation of the gross weights improves the estimation precision significantly only from wave 6 onwards, while the influence of the extreme weights in the preceding waves is comparatively small even without truncation.

Tabelle 19 Overview of the truncation of the weights

	threshold value for truncation	without truncation var(μ_W)/var(μ)	with truncation var(μ_W)/var(μ)	truncated weights number	in %
Wave 1	3	1.3556	1.1148	41	1.11
Wave 2	2	1.3210	1.0833	40	1.20
Wave 3	3	1.3027	1.1407	39	1.19
Wave 4	3	1.3400	1.1949	40	1.79
Wave 5	3.5	1.4315	1.3289	42	1.88
Wave 6	4	2.0066	1.5686	43	2.09
Wave 7	5	4.5207	2.0487	45	2.30
Wave 8	5	9.2769	2.4497	50	2.49

Note: Coefficients refer to the sample realized per wave.

4.3 Overview of the weighting variables

A total of three weights are available per survey wave. The first of these weights is the wave-specific gross weight (t**X**weight). In addition, a calibrated and untruncated weight (t**X**weight_k) and a truncated and calibrated weight (t**X**weight_g) are included in the data.[14] It should be noted that the upper benchmarks shown in Table 19 differ from the maximum values of the weights in Table 20. Since the wave-specific weights must have a mean value of 1 in relation to the sample realized within each shaft, the weights were recalibrated to the mean value of 1 after the truncation.

[13] The key points of the cutting refer to the gross weights Gti according to equation 1, which were recalibrated to an average value of 1. Since there is a margin of discretion in determining the basic parameters of the land use, the DAB data also include the untruncated gross weights, which allows the land use to be determined individually as required.

[14] In earlier publications of DAB data, extrapolation weights were also provided. If you are interested in extrapolation weights please contact the DAB project team: dab@edu.unibe.ch.

Tabelle 20 Overview of the inferential statistical weighting variables, waves 1 to 8

	mean	sd	min	max	N
W1: Raw weight (t1weight)	15.349	9.070	8.226	93.293	4083
W1: Weight calibrated (t1weight_k)	1.000	0.596	0.539	6.118	3680
W1: Weight truncated (t1weight_g)	1.000	0.339	0.559	3.108	3680
W2: Raw weight (t2weight)	18.715	10.259	9.128	104.954	4060
W2: Weight calibrated (t2weight_k)	1.000	0.567	0.488	5.606	3331
W2: Weight truncated (t2weight_g)	1.000	0.289	0.510	2.091	3331
W3: Raw weight (t3weight)	21.592	11.732	9.424	149.976	3735
W3: Weight calibrated (t3weight_k)	1.000	0.550	0.441	6.414	3281
W3: Weight truncated (t3weight_g)	1.000	0.375	0.453	3.083	3281
W4: Raw weight (t4weight)	31.463	18.040	11.464	229.546	2655
W4: Weight calibrated (t4weight_k)	1.000	0.583	0.368	6.668	2236
W4: Weight truncated (t4weight_g)	1.000	0.441	0.378	3.082	2236
W5: Raw weight (t5weight)	43.608	31.699	11.056	335.395	2864
W5: Weight calibrated (t5weight_k)	1.000	0.657	0.282	8.295	2229
W5: Weight truncated (t5weight_g)	1.000	0.574	0.287	3.558	2229
W6: Raw weight (t6weight)	83.337	139.739	12.750	2422.615	2738
W6: Weight calibrated (t6weight_k)	1.000	1.007	0.238	12.755	2061
W6: Weight truncated (t6weight_g)	1.000	0.686	0.255	4.265	2061
W7: Raw weight (t7weight)	109.211	202.216	13.992	3285.611	2496
W7: Weight calibrated (t7weight_k)	1.000	1.876	0.182	42.665	1958
W7: Weight truncated (t7weight_g)	1.000	1.024	0.205	5.645	1958
W8: Raw weight (t8weight)	158.243	412.973	14.521	7824.032	2400
W8: Weight calibrated (t8weight_k)	1.000	2.876	0.135	73.167	2009
W8: Weight truncated (t8weight_g)	1.000	1.204	0.167	6.161	2009

4.4 Use of the weighting variables

Generally, the weight of the survey wave from which variables are included in the analyses is to be used. If, for example, when data up to and including wave 3 are used in the analysis, the weight of wave 3 is to be used.

Statistics programs differ in the handling of weighting variables. In Stata, the wave-specific weighting variables can be used in different ways.[15] On the one hand, the weights can be specified within the scope of the available commands. In the documentation of the commands, it is explained in each case which types of weights can be used.[16] On the other hand, the command prefix `svy` can be used, which requires that the variables for identifying the survey design and weighting variables are specified in advance with the `svyset` command.

[15] These are suggestions, each of which must be specified for the respective analysis purpose and used sample.
[16] See in general: help weights.

```
. fre t1birthch [aweight=t1weight_g]
```

t1birthch ── Geburtsland: Schweiz

		Freq.	Percent	Valid	Cum.
Valid	1 ja	3341.121	90.79	90.79	90.79
	2 nein	338.8791	9.21	9.21	100.00
	Total	3680	100.00	100.00	

```
.
. svyset psu [pweight=t1weight_g], strata(stratum) fpc(fpc)

      pweight: t1weight_g
          VCE: linearized
  Single unit: missing
     Strata 1: stratum
         SU 1: psu
        FPC 1: fpc

.
. svy: proportion t1birthch
(running proportion on estimation sample)
```

Survey: Proportion estimation

Number of strata	= 31	Number of obs	= 3,680
Number of PSUs	= 223	Population size	= 3,680
		Design df	= 192

		Linearized	Logit	
	Proportion	Std. Err.	[95% Conf.	Interval]
t1birthch				
ja	.9079133	.0009961	.9059297	.9098592
nein	.0920867	.0009961	.0901408	.0940703

5 Data structure

The DAB data is provided in wave-specific data sets. The parent and students surveys of the first survey wave were published in individual data sets and a data file with additional cross-wave data (identification, weighting variables, socio-demographic characteristics and school information) is also published.

All DAB panel respondents are assigned an individual ID (variable: code), which is identical for all waves. The individual ID can be used to link the data of the waves required for each analysis. In general, the variable names start with the prefix t**X** where **X** stands for the respective wave. The variables of the parent survey are marked with the prefix **e**. The *first three survey* waves focus on socio-demographic information, educational decisions and subjective cost and benefit assessments. In order to obtain detailed information on the desired upper secondary education, the respondents are filtered according to a general part of the questionnaire as follows:

> *Depending on whether people are aiming for vocational apprenticeship following 9th grade (**filter 1, 2, 3**), school education (**filter 4 / 5**), an interim solution (**filter 6**), direct entry into employment (**filter 8**), are still undecided (**filter 7**) or have other plans (**filter 9**).*

The variable names of the filters of waves 1 to 3 are structured as follows:

> t **X f** variable name

with:

> **X**: Wave 1–3
>
> **f**: Filter 1–9

The survey *waves four to nine* focus on the documenting of educational and occupational trajectories. The respondnets are asked whether the last activity reported is still current and, if not, which activity(ies) have been performed since the last activity was terminated. The occupational and educational situations are divided into eight categories. With the help of a detailed filter guide, various characteristics of the training and activities are recorded:

> ***Filter A**: employment; **Filter B**: further training; **Filter C**: VET; **Filter D**: university study; **Filter E**: school based education; **Filter F**: internship; **Filter G**: military and civilian service; **Filter H**: unemployment; **Filter Z**: non-assignable activity. In addition, from the sixth wave onwards, parallel activities were also surveyed, for example, if a continuing VET course was attended part-time.*

The episode structure is reflected in the last digit of the variable name and documents the individual training or professional episodes that have been completed during the observation period of a specific wave. The first episode is always the one that was exercised at the time of the previous survey. If this first episode was completed during the time between surveys and a new activity was started, this is marked by a 2 and so on. The variable naming of the activity variables from wave 4 onwards is structured as follows:

t **X** f**Y** variable name **Z**

with:

X: Wave 4–8

Y: Filter A–Z

Z: Episode number: 1–6

The variable `t5fEstartm3`, for example, was collected in the fifth wave (t5) and contains the starting month (startm) of school education (fE), which represents the third activity episode (3) of the respondent since the time of the survey in the fourth wave.

Furthermore, waves 5 and 6 contain a detailed query of future plans following the current activity, which in turn was subject to a filter structure.

***Filter a**: vocational further training; **Filter b**: employment; **Filter c**: university study; **Filter d**: VET; **Filter e**: school based education; **Filter f**: internship; **Filter g**: Interim solution*

For example, the variable `t6fetyp` was collected in the sixth wave (t6) and contains the type of school education (fetyp) that is aimed for after the end of the current job.

In addition to the wave-specific codebooks contained in the data, a cross-wave codebook of the data from the DAB panel study provides an overview of which variables were (recurrently) collected in which survey waves. The wave-specific codebooks contain detailed information on the concepts and scales covered in each wave.

6 Missings – Coding of missing values

In the data of the DAB panel study, different types of missing values are reported as follows:

«**.**» = Person did not participate in the interview in wave X.

«**.l**» = Left censored person in waves 1 and 2, i.e. person had not yet entered the class at the time of the survey.

«**.m**» = Person participated in interview, but did not answer the question.

«**.n**» = Person participated in survey, but the question was not submitted due to the filter guidance.

«**.p**» = Person has filled out the questionnaire in paper format. Paper questionnaires were offered in a shortened version, which is why not all questions were presented to these people.

«**.t**» = Person has left class after wave 1 or 2 and in wave 2 and/or 3 filled out the questionnaire individually rather than in the classroom survey. The questionnaires for individual participation were offered in a shortened version, which is why not all questions were presented to these people.

7 Availability of data and citation

The data from the eight seven survey waves are freely accessible at FORS as public-use files.

FORS Center
c/o Université de Lausanne
Bâtiment Géopolis
CH-1015 Lausanne

Ref Projekt: 10773
https://swissubase.ch/en/catalogue/studies/10773/14509/overview

Bibliografische Zitierung der Daten:
DAB panel study(DAB): W1 - W8 [Dataset]. Universität Bern, Institut für Erziehungswissenschaft, Abteilung Bildungssoziologie. Distributed by FORS, Lausanne. https://doi.org/10.23662/FORS-DS-XXX-X

8 Contact

Project homepage www.dab.edu.unibe.ch

E-Mail dab@edu.unibe.ch

Telephone +41 31 631 53 56

Universität Bern
Institut für Erziehungswissenschaft
Abteilung Bildungssoziologie
Fabrikstrasse 8
CH- 3012 Bern

Literatur

Arrow, Kenneth (1973). «The Theory of Discrimination». In: *Discrimination in Labor Markets*. Hrsg. von Orley Ashenfelter und Albert Rees. Princeton: Princeton University Press, S. 193–216.

Becker, Gary S (1964). *Human Capital. A Theoretical and Empirical Analysis with Special Reference to Education*. New York: Columbia University Press.

Becker, Rolf (2000). «Klassenlage und Bildungsentscheidungen. Eine empirische Anwendung der Wert-Erwartungstheorie.» In: *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 52.3, S. 450–474. ISSN: 0023-2653.

— (2003). «Educational Expansion and Persistent Inequalities of Education: Utilizing Subjective Expected Utility Theory to Explain Increasing Participation Rates in Upper Secondary School in the Federal Republic of Germany». In: *European Sociological Review* 19.1, S. 1–24. ISSN: 1468-2672.

— (2012a). «Bildungsungleichheit im Lichte aktueller Theorieanwendung in der soziologischen Bildungsforschung». In: *Bildung–Arbeit–Erwachsenwerden. Ein interdisziplinärer Blick auf die Transition im Jugend- und jungen Erwachsenenalter*. Hrsg. von Max Bergmann, Sandra Hupka-Brunner, Thomas Meyer und Robin Samuel. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 43–75. DOI: 10.1007/978-3-531-19071-6_3.

— (2012b). «Der Übergang ins Hochschulstudium: Prozesse und Mechanismen am Beispiel der deutschen Schweiz». In: *Bildung–Arbeit–Erwachsenwerden. Ein interdisziplinärer Blick auf die Transition im Jugend- und jungen Erwachsenenalter*. Hrsg. von Max Bergmann, Sandra Hupka-Brunner, Thomas Meyer und Robin Samuel. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 305–331. DOI: 10.1007/978-3-531-19071-6_15.

Becker, Rolf und David Glauser (2018). «Are Prepaid Monetary Incentives Sufficient for Reducing Panel Attrition and Optimizing the Response Rate? An Experiment in the Context of a Multi-Wave Panel with a Sequential Mixed-Mode Design». In: *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique* 137.1, S. 1–22. ISSN: 0759-1063. DOI: 10.1177/0759106318762456.

Becker, Rolf, Sara Möser und David Glauser (2019). «Cash vs. vouchers vs. gifts in web surveys of a mature panel study––Main effects in a long-term incentives experiment across three panel waves». In: *Social Science Research*. DOI: https://doi.org/10.1016/j.ssresearch.2019.02.008.

Boudon, Raymond (1974). *Education, Opportunity, and Social Inequality: Changing Prospects in Western Society*. New York: Wiley.

Breen, Richard und John H Goldthorpe (1997). «Explaining Educational Differentials: Towards a Formal Rational Action Theory». In: *Rationality and Society* 9.3, S. 275–305. ISSN: 1043-4631.

Dillman, Don A (2000). *Mail and Internet Suveys. The Tailored Design Method*. New York: Wiley.

Dillman, Don A, Jolene D Smyth und Leah Melani (2014). *Internet, Mail, and Mixed-Mode Surveys: The Tailored Design Method*. Hoboken: John Wiley und Sons.

EDK, [Schweizerische Konferenz der Erziehungsdirektoren] (2014-09-19 2011). *Klassengrösse, EDK/IDES-Kantonsumfrage 2010-2011*. [Schweizerische Konferenz der kantonalen Erziehungsdirektoren]. URL: http://www.edk.ch/dyn/15293.php (besucht am 19. 09. 2014).

Erikson, Robert und Jan O Jonsson (1996). «Explaining Class Inequality in Education: The Swedish Test Case». In: *Can Education be Equalized*. Hrsg. von Robert Erikson und Jan O Jonsson. Boulder: Westview Press, S. 1–63.

Esser, Hartmut (1999). *Soziologie: Spezielle Grundlagen, Band 1: Situationslogik und Handeln*. Frankfurt a. M.: Campus Verlag. ISBN: 3593371448.

Glauser, David (2015). *Berufsausbildung oder Allgemeinbildung. Soziale Ungleichheiten beim Übergang in die Sekundarstufe II in der Schweiz*. Wiesbaden: Springer VS. DOI: 10.1007/978-3-658-09096-8_2.

Keller, Suzanne und Marisa Zavalloni (1964). «Ambition and Social Class: A Respecification». In: *Social Forces* 43.1, S. 58–70. ISSN: 1534-7605.

Kish, Leslie (1992). «Weighting for Unequal Pi». In: *Journal of Official Statistics* 8.2, S. 183–200.

Mincer, Jacob (1974). *Schooling, Experience, and Earnings*. New York: NEBR.

Sacchi, Stefan (2011). *TREE-Längsschnittgewichtung. Konstruktion und Anwendung. Dokumentation zu den Erhebungswellen 2000 bis 2010*. Basel/Zürich: TREE und cue sozialforschung.

Sewell, William H, Archie O Haller und Murray A Straus (1957). «Social Status and Educational and Occupational Aspiration». In: *American Sociological Review* 22.1, S. 67–73. ISSN: 0003-1224.

Thurow, Lester C (1975). *Generating Inequality: Mechanisms of Distribution in the U.S. Economy*. New York: Basic Books.