Research Data

Tobias Koberg

Anonymization Procedures

NEPS Starting Cohort 2— Kindergarten
From Kindergarten to Elementary School

Scientific Use File Version 8.0.1

1 Preamble

This documentation gives an exhaustive explanation of all disclosure risk minimization techniques applied before dissemination to the Scientific Use File (SUF) of Starting Cohort (SC) 2, version 8.0.1 (doi:10.5157/NEPS:SC2:8.0.1). For a quick reference to all variables and datafiles which are affected in detail and on which level you will find the full information, please skip forward to appendix section A.1. If you have inquiries or need further information not covered in this document please feel free to contact the research data center at fdz@lifbi.de. Please note the following general remarks:

- in all tables, the anonymization suffix (_O, _R, _D) was removed from the variable name to enhance readability and remove duplicates. To find a specific variable in the data file, add the corresponding suffix or use a wildcard.

- to ensure a data-consistent documentation, all tables in this document have been produced on basis of the SUF data. Variable labels in the data have been cut off at 80 characters, which inherits to this documentation.

- as additional documentation, a *Semantic Data Structure File* is available. This is the complete SUF where all cases have been removed. Feel free to use this as a first overview to the data in cases you do not have access to the real data (e.g. you are interested in data files which are available OnSite only). You are well advised to use this preparing your syntax *before* initiating your travel to our OnSite facility. You find it freely available on the documentation page to this SUF:

  ```
  www.neps-data.de > Data Center > Data and Documentation > Starting Cohort
  Kindergarten > Documentation
  ```

1.1 Specifications

To ensure the best possible confidentiality protection of individuals and individual micro data, the National Educational Panel Study complies with strict international standards. To achieve this, a *portfolio approach*, based on the ideas of **LaneHeusMulcahy2008.20130315**, has been established. This approach interlocks data security mechanism on five different levels, while the actual data modification (i.e. anonymization) is one of them. For further information, see the article by **koberg2016disclosing**. Operationalizing this modification procedure, it can be abstracted to the following two criteria:

- the disseminated data was transferred to so called *de facto anonymous data*. Identifiable information is coarsened or cut off and kept securely to minimize the risk of statistical disclosure.

- the use of data is strictly confidential and for statistical purposes only. The closed contract only grants access to members of the scientific community. This contract has a vast amount of legal stipulations, one of them being a large fine which applies for the realization of re-identification on purpose. Therefore, the disseminated data is highly protected by law and allows a more flexible range of available data.

To pick up the latter, the NEPS has made a huge effort regarding legal regulations to offer as much analysis power of data as possible. This *paradigm of information esteem* reveals the fact that conducted measures of statistical disclosure control are few. Also, if there really was a need for modification, only non-perturbative methods were used.

1.2   Onion-shaped model

The NEPS grants the user three different modes of data access: (1) *OnSite*, which stands for the opportunity to use the secured infrastructure made available at the LIfBi in Bamberg, (2) *RemoteNEPS*, which is a progressive remote access technology providing a virtual desktop, and finally (3) *Download*, indicating the possibility to fetch data via a secure web portal.

These given access modes have been originated to allow anonymization routines for a subtle differentiation of information. The three resulting levels of anonymization define as follows:

- data provided OnSite is generally not further anonymized. However, even those data was rendered *de facto anonymous*, for no disclosure risk to persist. All information contained remains completely sane. Although users have to deal with limited possibilities of data access (i.e. supervised import and export of their results), they are free to work with all data available in a secure environment.

- access via RemoteNEPS is considered equivalent to OnSite, hence most of the data stays complete.

- as Download is assumed to be the most hazardous access mode[1], some more anonymization techniques are done to the dataset.

Obviously this approach results in three different versions of all involved datasets. To enable a consistent structure, these data files always contain the entire set of variables; it is their content which differs through the three levels.

As normally there is no need to resign aggregated variables in the higher levels (i.e. OnSite or RemoteNEPS), those are already defined as a surplus to the original variable in the OnSite-version. Stepping down to RemoteNEPS, the content of related variables too sensitive for this level is overwritten with an exclusive missing code – an operation which we define as *purging*. Note that system missing values are not affected, allowing the user to differ between value existence and nonexistence. This still is a valuable additional information. The same logic applies to the transition from RemoteNEPS to Download.

---

1   'hazardous' in terms of: the downloaded content is no longer under physical control of the LIfBi

While there is no explicit documentation to this fact, it should remain clear that this procedure accumulates, i.e. purged content under RemoteNEPS is therefore neither included in RemoteNEPS nor in Download.

This *onion-shaped model* provides both ease of (1) use of different sensitivity models (e.g. preparing an analysis using the Download dataset and conducting it afterwards using the OnSite-data) and (2) documentation, for the subject of documentation is the most sensitive level (OnSite), with RemoteNEPS and Download levels being a subset of these data.

The fourth layer *master* depicted below contains every material which is needed during data processing by the NEPS, but is not meant for the scientific community to be usable.
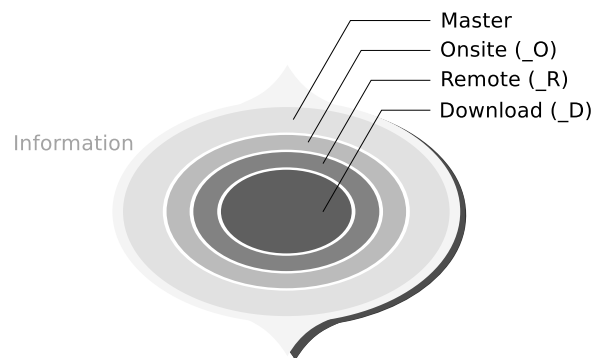

Figure 1: Onion-shaped model defining the different anonymization levels

Technically, this model realizes in a single letter suffixed to dataset and variable names. All datasets available OnSite only are marked with an additional **_O**, those available via RemoteNEPS with **_R** and Download files with **_D**. The same procedure applies when it comes to variable differentiation. A variable which is available unaltered OnSite only is suffixed with **_O**. In RemoteNEPS-access or Download, this variable is still present but purged. If there is an alternate version (e.g. with coarsened content) for RemoteNEPS (suffix **_R**) or Download (suffix **_D**), those can be used. As said before, these are already integrated in the OnSite version. See section 3 for some examples (especially for a practical explanation what is meant by *purging* here).

2 Conducted measures

Keeping the usability and the paradigm of information esteem in mind, only very few alterations are actually done to the dataset. Please note that all information is still available somewhere and that only RemoteNEPS and (mainly) the Download version are constraint in this matter. In fact, in this SUF, only 330 out of 4651 variables are modified in some way.

The following table gives an explanatory overview of all measures conducted. The remainder of this section gives a more detailed description of those topics.

Table 1: Availability of sensitive data

Topic	OnSite	RemoteNEPS	Download
String variables	anonymized	n/a	n/a
International[1]	full data	full data	collapsed
Regional (National)	restricted	restricted	restricted
Institutional context	full data	anonymized	n/a
Macro indicators	accessible	n/a	n/a
Others	full data	full data	aggregated

[1] international geographical information (e.g., nation states, national languages)

2.1 String variables

All variables containing actual text and whose type is therefore string are purged in the RemoteNEPS version. The information remains accessible OnSite. However, you should be aware that in the majority of cases, text entries have been recoded into derived variables (usually suffixed with _g1, _g2, etc.), which contain the information collected. Those variables are available in all three levels, only the actual text entry is available OnSite only. The text entries have been reviewed by staff to ensure that absolutely no reidentificational material is included. To indicate the availability of those variables OnSite only, the variable names have been suffixed with an additional _0.

Table 2: All string variables; those are only available onsite

file	variable	label
pCourseClass	ec20020	My lessons have no connection to a course, I use an open concept, specifically:

(...)

Table 2: (continued)

file	variable	label
pEducator	e212819	Further training, other, text
pEducator	e212823	Group leader: further education, additional certification, text
pEducator	e21980b	Kindergarten grp leaqder:professional qualification,other qualification,text
pEducator	e219817	Group leader: extent of work - other, text
pEducator	ec03091	Working hours: Other text
pGroups	e217406	Other developmental disorders (open)
pInstitution	h212013	Kindergarten: participation in quality improvement measure, other measure, text)
pInstitution	h212307	Kindergarten: frequency of honorary tasks, other, text
pInstitution	h212819	Training, other, text
pInstitution	h217006	Kindergarten: number of children with other diagnosed developm. disorders, text
pInstitution	h218020	Kindergarten: collaboration with a school: other, text
pInstitution	h21905t	Kindergarten: problems with surroundings, other, text
pInstitution	h21920a	Kindergarten: diagnostic offers, other, text
pInstitution	h21920b	Kindergarten: therapeutic offers, other, text
pInstitution	h21920c	Kindergarten: offers for parents, other, text
pInstitution	h21932t	Other qualification, text
pInstitution	h21980a	Kindergarten management: professional qualification: other qualification: Text
pInstitution	h219817	Use of actual weekly working hours, other, text
pInstitution	h22802t	School: cooperation with kindergarten: other measures, text
pInstitution	h22803t	School: Cooperation with secondary schools: Other – text
pInstitution	h229003	School: approach, educational, text
pInstitution	h229005	School: approach, promotion, text
pInstitution	h229007	School: approach, integration, text
pInstitution	h229009	School: approach, other, text
pInstitution	h22903t	School: problems in the environment: other, text
pInstitution	h40184a	Measures for parents with an immigration background open 1
pInstitution	h40185a	Measures intercultural competence open for educators (open)
pInstitution	h416110	Remedial language teaching - open
pInstitution	h417120	Migration-specific support measures - open
pInstitution	h417160	Migration-specific additional support for teachers: Open specification
pInstitution	h418012	Language remediation: individual remediation - other, open
pInstitution	h418052	Language remediation: small group remediation - other, open
pInstitution	h418092	Language remediation: whole group remediation - other, open
pInstitution	hb1006t	Execution of the speech remediation persons with other qualification, text
pInstitution	hb1015e	Other from of bilingual care
pInstitution	hc1007d	Remedial language teaching: other type of assistance (open)
pInstitution	hd0041b	No entry made
pParent	p19020j	Special needs OPEN
pParent	p262605	Outside activities: Usage other, Text

(…)

Table 2: (continued)

file	variable	label
pParent	p401001	Migrant status Other
pParent	p403041	Migrant status partner - other
pParent	p524853	AD(H)D: Medications - open
pParent	p723090	Type of school (open)
pParent	p731803	Highest educational qualification Respondent, type open
pParent	p731814	Vocational qualification respondent (open)
pParent	p731817	Type tertiary qualification respondent (open)
pParent	p731853	Highest educational qualification Partner, type open
pParent	p731864	Professional qualification partner (open)
pParent	pc13140	Deferral of enrollment: attendance at current facility, other
pTarget	t27111d	Student: Courses outside of school: Other courses: Text, where
pTarget	t400071	Other country of birth of the mother
pTarget	t400091	Other country of birth of the father
pTarget	t410001	Language of origin of TP - open
pTarget	t435030	other religion or denomination
pTarget	t435050	Religion or denomination (other)
spParentGap	ps29102	Other activity
spParentSchool	p723090	Type of school (open)
spSibling	p723090	Type of school (open)

2.2 International: countries, languages, and nationalities

All information corresponding to international localization, nationality or languages is completely available only OnSite or via RemoteNEPS. Variables comprised in the Download SUF are aggregated to the binary information German/non-German. Please note that this variable does not necessarily has the same root of the variable name as the affected variable (e.g. if it is part of the instrument and not generated during the anonymization process). Look in the vicinity of the documented variable to find it. Also, see section 3.3.

Table 3: All variables which are affected by country/language/nationality aggregation

file	variable	label
pParent	p32601a	Position generator: country nurse or male nurse
pParent	p32601b	Position generator: country engineer
pParent	p32601c	Position generator: country warehouse/transport worker
pParent	p32601d	Position generator: country social worker
pParent	p32601e	Position generator: country sales clerk
pParent	p32601f	Position generator: country police officer
pParent	p32601g	Position generator: country doctor

(...)

Table 3: (continued)

file	variable	label
pParent	p32601h	Position generator: country banker
pParent	p32601k	Position generator: country car mechanic
pParent	p32601l	Position generator: country legal practitioner
pParent	p32601m	Position generator: country optician
pParent	p32601n	Position generator: country translator
pParent	p32601o	Position generator: country teacher at elementary school, Hauptschule or Real...
pParent	p400010_g1	Country of birth respondent
pParent	p400010_g2	Country of birth respondent (aggregated)
pParent	p400070_g1	Country of birth mother respondent
pParent	p400070_g2	Country of birth respondent's mother (aggregated)
pParent	p400090_g1	Respondent's father's country of birth
pParent	p400090_g2	Country of birth respondent's father (aggregated)
pParent	p400500_g3	group of origin
pParent	p400500_g3v1	Group of origin
pParent	p401150_g2	Nationality Respondent not German (aggregated)
pParent	p403010_g1	Country of birth partner abroad
pParent	p403010_g2	Country of birth partner (aggregated)
pParent	p403070_g1	Country of birth mother partner
pParent	p403070_g2	Country of birth partner's mother (aggregated)
pParent	p403090_g1	Country of birth father partner
pParent	p403090_g2	Country of birth partner's father (aggregated)
pParent	p404050_g2	Other nationality partner (aggregated)
pParent	p406010_g1	Country of birth target child
pParent	p406010_g2	Country of birth target child (aggregated)
pParent	p407050_g2	Nationality Target child (aggregated)
pParent	p407060_g2	Target child's second nationality (aggregated)
pParent	p731804_g1	Highest educational qualification of respondent abroad (country)
pParent	p731804_g2	Country of respondent's highest educational qualification (aggregated)
pParent	p731823_g1	Country of vocational qualification (additional response)
pParent	p731823_g2	Country of professional qualification (additional response, aggregated)
pParent	p731854_g1	Country of partner's highest school-leaving qualification
pParent	p731854_g2	Country of highest educational qualification partner (aggregated)
pParent	p731873_g1	Country of vocational qualification, partner, (additional response)
pParent	p731873_g2	Country of partner's professional qualification (add. response, aggregated)
pTarget	t400000_g1	Birth country
pTarget	t400000_g2	Country of birth (aggregated)

(...)

Table 3: (continued)

file	variable	label
spParentSchool	p723060_g1	Country of school
spParentSchool	p723060_g2	Country of school (aggregated)
ParentMethods	px80204	Interview: Interview language (realized case)
TargetMethods	tx80204	Interview: Interview language (realized case)
pEducator	e41100a_g2	Mother tongue (response 1, ISO 639.2)
pEducator	e41100a_g3	Mother tongue (response 2, ISO 639.2)
pEducator	e41100a_g4	Mother tongue (response 3, ISO 639.2)
pEducator	e41100a_g5	Mother tongue (response 4, ISO 639.2)
pParent	p410000_g1	First language/mother tongue child (ISO 639.2)
pParent	p410000_g2	First language/mother tongue child (aggregated)
pParent	p410002_g1	Further first language/ mother tongue child (ISO 639.2)
pParent	p410002_g2	Further first language/ mother tongue child (aggregated)
pParent	p412001_g1	Interaction language household, detailed (ISO 639.2)
pParent	p412001_g2	Interactive language household detailed (aggregated)
pParent	p413000_g1	First language/mother tongue respondent (ISO 639.2)
pParent	p413000_g2	First language/mother tongue respondent (aggregated)
pParent	p413002_g1	Further first language/ mother tongue respondent (ISO 639.2)
pParent	p413002_g2	Further first language/ mother tongue respondent (aggregated)
pParent	p414000_g1	First language/mother tongue partner (ISO 639.2)
pParent	p414000_g2	First language/mother tongue partner (aggregated)
pParent	p414002_g1	Further first language/ mother tongue partner (ISO 639.2)
pParent	p414002_g2	Further first language/ mother tongue partner (aggregated)
pParent	p414100_g1	First language/mother tongue new partner (ISO 639.2)
pParent	p414100_g2	First language/mother tongue new partner (aggregated)
pParent	p414102_g1	Further first language/ mother tongue new partner (ISO 639.2)
pParent	p414102_g2	Further first language/ mother tongue new partner (aggregated)
pTarget	t41000a_g2	Mother tongue (response 1, ISO 639.2)
pTarget	t41000a_g3	Mother tongue (response 2, ISO 639.2)
pTarget	t41000a_g4	Mother tongue (response 3, ISO 639.2)
pTarget	t41000a_g5	Mother tongue (response 4, ISO 639.2)
pParent	p401150_g1	Citizenship respondent not German

(...)

Table 3: (continued)

file	variable	label
pParent	p404050_g1	Other citizenship partner
pParent	p407050_g1	Citizenship of the target child
pParent	p407060_g1	Second citizenship of the target child

2.3   Regional localization

For all variables which register a regional position – such as places of birth, work, school, or residence – the following variables have been generated (if possible):

- an indicator for West Germany and East Germany (including Berlin)

- the federal state for this location.

As NEPS data which has been surveyed in institutional context enjoys special focus in this matter, federal states are only available in RemoteNEPS. When using data collected in connection with schools or higher education institutions, it is not allowed to use Federal-State-related information directly or indirectly contained in the data for analyses aiming at a direct Federal State comparison, direct conclusions to be drawn about a Federal State, or a reconstruction of the concrete Federal State affiliation of persons, households, and institutions. Please see the Data Use Agreement for further information about this topic. Regional localization below this level is not directly available. If you have need to merge data using a regional key below those levels, please contact the LIfBi RDC (fdz@lifbi.de), which may deal with this matter for you.

Table 4: All regional variables which are only available in RemoteNEPS

file	variable	label
CohortProfile	tx80109_g2	Sample: Federal state school
WeightsElementarySchool	stratum_imp2	implicit stratum 2 (federal states of institution)
WeightsElementarySchool	stratum_imp3	implicit stratum 3 (regional classification of institution)
WeightsElementarySchool	stratum_imp4	implicit stratum 4 (organizing institution)
pParent	p723050	Federal state of the school
pParent	p751001_g2	Place of Residence (Federal State)
pParent	pb11610_g2	Place of elementary school (federal state)
spParentSchool	p723030_g2	Municipality of school (Federal state)
spSibling	p723050	State of school

2.4 Institutional characteristics

In this Starting Cohort, also information about institutions is available. A special focus of anonymization was directed to protection of institutional data, i.e. information about kindergarten, schools, and universities, but also educators and teachers. To minimize the risk of re-identification, this data is only available in RemoteNEPS. This includes the complete datafile(s)

```
pInstitution,
```

but may also affect basic structural details about the group or class (e.g. size of class) in other datafiles. Furthermore, personal information about educators and teachers is treated more securely. You will find detailed information about these subjects from RemoteNEPS onward. Refer to section 2.6 to see the affected variables.

2.5 Macro indicators

Additional information including structural topography and macroeconomic measures was collected. You find those available OnSite only. The following data is available:

- Microm data, describing the surrounding of the targets residence

- Microm data, describing the surrounding of the school

- RegioInfas (infas geodaten) data, describing the surrounding of the targets residence

- RegioInfas data, describing the surrounding of the school

2.6 Other variables

All other variables affected by anonymization and not classifiable in the previous groups are listed below. In the first table (table 5), you find all variables which were aggregated in the Download-SUF. An additional table below each entry displays the values of the variable, both as found in RemoteNEPS (if table column is empty, the variable is numeric and does not have a value label scheme) and in Download.

The second table (table 6) contains all variables which are not aggregated, but purged completely in the Download-version, that is, there is no aggregated version.

2 Conducted measures

Table 5: Other variables which were aggregated in the Download-SUF (aggregation is also stated)

[file] `variable: label`

`[CohortProfile]` tx80505: List of children / students: Special educational needs Target child (aggregated)

RemoteNEPS		Download
No special educational needs	**1**	No special educational needs
Special needs with focus reading	**2**	Special educational needs
Special needs with focus language	**3**	
Special needs with focus emotional and social development	**4**	
Special needs with focus autism	**5**	
Special needs with other priority	**6**	
unknown	**7**	
Special needs with focus seing	**8**	
Special needs with focus hearing	**9**	
Special needs with focus physical motoric development	**10**	
Special needs with focus mental development	**11**	
Special needs with focus learning/language	**23**	
Special needs with focus learning/emotional development	**24**	
Special needs with focus learning/autism	**25**	
Special needs with focus learning/other	**26**	
Special needs with focus language/emotional social development	**34**	
Special needs with focus language/others	**36**	
Special needs with focus emotional social development/autism	**45**	
Special needs with focus emotional social development/other	**46**	
Special needs with focus autism/seing	**58**	
by SK marked with cross/line	**96**	
not administered	**97**	
No indication	**99**	
Special needs with focus learning/language/emotional social development	**234**	
Special needs with focus Learning/language/Autism	**235**	
Special needs with focus learning/language/Others	**236**	
Special needs with focus Language/emotional social development/others	**346**	
Special needs with focus learning/language/mental development	**2311**	
Special needs with focus learning/emotional social mental development	**2411**	

`[pCourseClass]` e227400_g1: Class: number of students, total (simplified)

RemoteNEPS		Download
	1	Below 10
	2	10 to 14
	3	15 to 19
	4	20 to 24
	5	25 to 29
	6	30 to 34
	7	35 and more

(...)

Table 5: (continued)

[pCourseClass] e229400: Class: Facilities: Classroom size (aggregated)

RemoteNEPS		Download
	1	under 30
	2	30 to 39
	3	40 to 49
	4	50 to 59
	5	60 and more

[pCourseClass] ec0150a: Number of lessons/week: German/literacy (aggregated)

RemoteNEPS		Download
	1	less than 5
	2	5 to 6
	3	more than 6

[pCourseClass] ec0151a: Number of lessons/week: Math (aggregated)

RemoteNEPS		Download
	1	less than 5
	2	5 to 6
	3	more than 6

[pCourseClass] ec0152a: Number of lessons/week: General knowledge/local history and geography (aggreg.)

RemoteNEPS		Download
	1	less than 2
	2	2 to 3
	3	more than 3

[pCourseClass] ec01530: Number of lessons/week of class in total (aggregated)

RemoteNEPS		Download
	1	0 to 20
	2	21
	3	22
	4	23
	5	24
	6	25 and more

(...)

Table 5: (continued)

[pEducator] e219800: Nursery school group leader: Professional qualification (categorized)

RemoteNEPS		Download
Educator	1	Nursery school teacher
Childcare worker	2	Other
Remedial teacher (Fachschule)	3	
Social education worker or social worker (Fach-hochschule or similar qualification)	4	
Qualified educator, qualified educationalist (university or comparable qualification)	5	
Remedial teacher (Fachhochschule or similar qualification)	6	
Intern in the probationary year	7	
Without completed vocational qualification	8	
Other qualification, namely:	9	

[pEducator] e219820: Group leader: duration professional activity: previous facilities (categorized)

RemoteNEPS		Download
	1	0-4
	2	5-9
	3	10-14
	4	15-19
	5	20-24
	6	25-29
	7	30-34
	8	35-39
	9	40 and more

[pEducator] e219821: Group leader: duration professional activity: current facility (categorized)

RemoteNEPS		Download
	1	0 to less than 5
	2	5 to less than 10
	3	10 to less than 15
	4	15 to less than 20
	5	20 to less than 25
	6	25 to less than 30
	7	30 to less than 35
	8	35 and more

[pEducator] e76212y: Year of birth (categorized)

RemoteNEPS		Download
	1	Before 1950
	2	1950-1959
	3	1960-1969
	4	1970-1979
	5	after 1979

(...)

Table 5: (continued)

[pGroups] e219410: Kindergarten core group: Room use: Number of rooms (categorized)

RemoteNEPS	Download	
	5	5 and more

[pGroups] e219411: Nursery school core group: Room use: Size of rooms (categorized)

RemoteNEPS	Download	
	1	0-19
	2	20-29
	3	30-39
	4	40-49
	5	50-59
	6	60-69
	7	70-79
	8	80-89
	9	90-99
	10	100 and more

[pParent] p731911: Number of employees respondent (aggregated)

RemoteNEPS		Download	
1 to less than 5	1	1 to less than 5	1
5 to less than 10	2	5 to less than 10	2
10 to less than 20	3	10 to less than 20	3
20 to less than 50	4	20 and more	4
50 to less than 100	5		
100 to less than 200	6		
200 to less than 250	7		
250 to less than 500	8		
500 to less than 1,000	9		
1,000 to less than 2,000	10		
2,000 and more	11		

[pParent] p731961: Number of employees partner (categorized)

RemoteNEPS		Download	
1 to less than 5	1	1 to less than 5	1
5 to less than 10	2	5 to less than 10	2
10 to less than 20	3	10 to less than 20	3
20 to less than 50	4	20 and more	4
50 to less than 100	5		
100 to less than 200	6		
200 to less than 250	7		
250 to less than 500	8		
500 to less than 1,000	9		
1,000 to less than 2,000	10		
2,000 and more	11		

(...)

Table 5: (continued)

[file] `variable: label`

`[spChildCare] fee: Care: Fees (Euro, categorized)`

RemoteNEPS		Download
	1	under 50
	2	from 50 to exclusively 100
	3	from 100 to exclusively 150
	4	from 150 to exclusively 200
	5	from 200 to exclusively 250
	6	from 250 to exclusively 300
	7	from 300 to exclusively 350
	8	from 350 to exclusively 400
	9	from 400 to exclusively 450
	10	from 450 to exclusively 500
	11	500 and more

Table 6: Other variables which were purged in the Download-SUF

file	variable	label
WeightsElementarySchool	tx80112	Amount of students in grade 2 total (accord. to the Federal Statistical Office)
pCourseClass	e19001c	Number students with special needs
pCourseClass	e199001	Number special needs classes: learning
pCourseClass	e199002	Number special needs classes: language
pCourseClass	e199003	Number special needs classes: physical/motor development
pCourseClass	e199004	Number special needs classes: emotional/social development
pCourseClass	e199005	Number special needs classes: mental development
pCourseClass	e199006	Number special needs classes: seeing
pCourseClass	e199007	Number special needs classes: hearing
pCourseClass	e199008	Number special needs classes: autism
pCourseClass	e227400	Class: number of female students
pCourseClass	e227401	Class: number of male students
pCourseClass	e451000	Daycare: number of students with migrant background
pCourseClass	e79201a	Class: share class affiliation - low class
pCourseClass	e79201b	Class: share class affiliation - middle class
pCourseClass	e79201c	Class: share class affiliation - upper class
pCourseClass	e79202a	Class: share parents with a higher education degree

(...)

Table 6: (continued)

file	variable	label
pEducator	e536020	Age Career choice
pEducator	e53702y	Year of university entrance qualification
pEducator	e537030	Federal state higher education entrance qualification (Federal state)
pEducator	e53710y	Start Teaching degree course
pEducator	e537150	Year of state examination
pEducator	e537200	Federal state second Staatsexamen (Federal state)
pEducator	e76212m	Month of birth
pGroups	e217401	Kindergarten core group: girls registered
pGroups	e217402	Kindergarten core group: boys registered
pGroups	e217403	Core group: children with diagnosed speech disorders (number)
pGroups	e217404	Core group: children with diagnosed behavioral disorders (number)
pGroups	e217405	Children with other developmental disorders (number)
pGroups	e217411	Core group: year of birth 2010 or later; number of children altogether
pGroups	e217412_w1	Core group: year of birth 2009 and later; number of children altogether
pGroups	e217412_w2	Core group: year of birth 2009; number of children, altogether
pGroups	e217413	Core group: year of birth 2008; number of children, altogether
pGroups	e217414	Core group: year of birth 2007; number of children altogether
pGroups	e217415	Core group: year of birth 2006; number of children altogether
pGroups	e217416_w1	Core group: year of birth 2005; number of children, altogether
pGroups	e217416_w2	Core group: year of birth 2005 or earlier: number of children altogether
pGroups	e217417	Core group: year of birth 2004 and earlier; number of children, altogether
pGroups	e217421	Core group: year of birth 2010 or later; number children attending up to 5 hours
pGroups	e217422_w1	Core group: year of birth 2009 and later; number children, attend up to 5 hours

(...)

Table 6: (continued)

file	variable	label
pGroups	e217422_w2	Core group: year of birth 2009; number of children attending up to 5 hours
pGroups	e217423	Core group: year of birth 2008; number of children, attending up to 5 hours
pGroups	e217424	Core group: year of birth 2007; number of children, attending up to 5 hours
pGroups	e217425	Core group: year of birth 2006; number of children attending up to 5 hours
pGroups	e217426_w1	Core group: year of birth 2005; number of children, attend up to 5 hours
pGroups	e217426_w2	Core group: year of birth 2005 or earlier: number children, attending up to 5h
pGroups	e217427	Core group: year of birth 2004 and earlier; number of children, attend up to 5 h
pGroups	e217431	Core group: year of birth 2010 or later; number of children attending 5-7 hours
pGroups	e217432_w1	Core group: year of birth 2009 and later; number of children, attend 5-7 hours
pGroups	e217432_w2	Core group: year of birth 2009; number of children attending 5-7 hours
pGroups	e217433	Core group: year of birth 2008; number of children, attending 5-7 hours
pGroups	e217434	Core group: year of birth 2007; number of children, attending 5-7 hours
pGroups	e217435	Core group: year of birth 2006; number of children, attending 5-7 hours
pGroups	e217436_w1	Core group: year of birth 2005; number of children, attend 5-7 hours
pGroups	e217436_w2	Core group: year of birth 2005 or earlier; number children, attending 5-7 hours
pGroups	e217437	Core group: year of birth 2004 and earlier; number of children, attend 5-7 hours
pGroups	e217441	Core group: year of birth 2010 or later; number children attending more than 7h
pGroups	e217442_w1	Core group: year of birth 2009 and later; nr. children, attend more than 7 hours
pGroups	e217442_w2	Core group: year of birth 2009; number of children attending more than 7 hours

(...)

Table 6: (continued)

file	variable	label
pGroups	e217443	Core group: year of birth 2008; number of children, attending more than 7 hours
pGroups	e217444	Core group: year of birth 2007; number of children, attending more than 7 hours
pGroups	e217445	Core group: year of birth 2006; number of children, attending more than 7 hours
pGroups	e217446_w1	Core group: year of birth 2005; number of children, attend more than 7 hours
pGroups	e217446_w2	Core group: year of birth 2005 or earlier; nr. children, attending more than 7h
pGroups	e217447	Core group: Year of birth 2004 and earlier; nr. children, attend over 7 hours
pGroups	e217451	Core group: year of birth 2010 and later; number of children, disability
pGroups	e217452_w1	Core group: year of birth 2009 and later; number of children, disability
pGroups	e217452_w2	Core group: year of birth 2009; number of children, disability
pGroups	e217453	Core group: year of birth 2008; number of children, disability
pGroups	e217454	Core group: year of birth 2007; number of children, disability
pGroups	e217455	Core group: year of birth 2006; number of children, disability
pGroups	e217456_w1	Core group: year of birth 2005; number of children, disability
pGroups	e217456_w2	Core group: year of birth 2005 or earlier; number of children, disability
pGroups	e217457	Core group: year of birth 2004 and earlier; number of children, disability
pGroups	e401200	Number of children with another interactive language
pGroups	e45110a	Core group: year of birth 2004 and earlier; number of children, mig. background
pGroups	e45110b_w1	Core group: year of birth 2005; number of children, migration background
pGroups	e45110b_w2	Core group: year of birth 2005 or earlier; number children, migrant background

(...)

Table 6: (continued)

file	variable	label
pGroups	e45110c	Core group: year of birth 2006; number of children, migrant background
pGroups	e45110d	Core group: year of birth 2007; number of children, migrant background
pGroups	e45110e	Core group: year of birth 2008; number of children, migrant background
pGroups	e45110f_w1	Core group: year of birth 2009 and later; number children, migration background
pGroups	e45110f_w2	Core group: year of birth 2009; number of children with migrant background
pGroups	e45110g	Core group: year of birth 2010 or later; number children with migrant background
pGroups	e79101a	Group: proportion class affiliation - low class
pGroups	e79101b	Group: proportion class affiliation - middle class
pGroups	e79101c	Group: proportion class affiliation - upper class
pGroups	e79102a	Group: proportion university graduates
pGroups	eb1000a	Language skills above average age - all children
pGroups	eb1000b	Language skills above average age-thereof children with migration background*
pGroups	eb1000c	Language skills average - all children
pGroups	eb1000d	Language skills average - portion of children with migrant background*
pGroups	eb1000e	Language skills below average - all children
pGroups	eb1000f	Language skills below average - portion of children with migrant background*
pInstitution	h76512m	Month of birth
pParent	p435030_g1	Religion and religiousness: Other religious community
pTarget	t400070	Country of birth of the mother
pTarget	t400090	Country of birth of the father
pTarget	t435030_g1	Religion and religiousness: other religious community
pTarget	t723400	grade attended

3 Examples

The following examples compare variables from some selected topics. On the left side, you see these variables as they can be found in the RemoteNEPS-SUF (i.e., this is the variant of the variable which contains the full information). On the right side, you find the same variables, but from the Download-version of the SUF, that is, the aggregated or purged variant of the variable. This illustrates the anonymization applied.

3 Examples

3.1 Number of employees

The first example is a variable collected during the interview of Starting Cohort 6 (Adults). The question further investigates the occupational periphery of employers, while here, this employer is the mother of the respondent. On the left, you see the complete information as found in the RemoteNEPS-SUF. This variable is collected using seven classes, where the last four are sparsely filled (variable t731406_R, table on the lower left). Because of this, those four codes have been recoded into one category (»*20 and more*«) and a new variable is generated (t731406_D, upper left). In the Download-SUF, you find this aggregated variable unaltered (upper right). The variable t731406_R rich in content was purged, though. You see that all meaningful codes have been recoded to »-*53 Anonymized*«. Note that the missing code »-*54 Missing by design*« is transported to the Download-SUF unmodified, so you already can derive the actual *statistical power* of this variable.

RemoteNEPS	Download

```
SC6_pTarget_R_5-1-0.dta
(NEPS SUF, SC6 5.1.0 (remote); doi:10.5157/NEPS:SC6:5.1.0)

t731406_D -- Number of mother's employees (categorized)
------------------------------------------
                          |    Freq.
--------------------------+---------------
Valid  -98 Don't know     |         7
       -97 Refused        |         1
       -54 Missing by design |     36700
       0   none           |       423
       1   1 to 4         |       330
       2   5 to 9         |        64
       3   10 to 19       |        22
       4   20 and more    |        28
           Total          |     37575
Missing .                 |     15982
Total                     |     53557
------------------------------------------


t731406_R -- Number of mother's employees
------------------------------------------
                          |    Freq.
--------------------------+---------------
Valid  -98 Don't know     |         7
       -97 Refused        |         1
       -54 Missing by design |     36700
       0   None           |       423
       1   1 to 4         |       330
       2   5 to 9         |        64
       3   10 to 19       |        22
       4   20 to 49       |        21
       5   50 to 99       |         3
       6   100 to 199     |         3
       7   200 to 249     |         1
           Total          |     37575
Missing .                 |     15982
Total                     |     53557
------------------------------------------
```

```
SC6_pTarget_D_5-1-0.dta
(NEPS SUF, SC6 5.1.0 (download); doi:10.5157/NEPS:SC6:5.1.0)

t731406_D -- Number of mother's employees (categorized)
------------------------------------------
                          |    Freq.
--------------------------+---------------
Valid  -98 Don't know     |         7
       -97 Refused        |         1
       -54 Missing by design |     36700
       0   none           |       423
       1   1 to 4         |       330
       2   5 to 9         |        64
       3   10 to 19       |        22
       4   20 and more    |        28
           Total          |     37575
Missing .                 |     15982
Total                     |     53557
------------------------------------------


t731406_R -- Number of mother's employees
------------------------------------------
                          |    Freq.
--------------------------+---------------
Valid  -54 Missing by design |     36700
       -53 Anonymized     |       875
           Total          |     37575
Missing .                 |     15982
Total                     |     53557
------------------------------------------
```

3 Examples

3.2 Countries (aggregated list)

To illustrate the anonymization of countries, first have a look at this variable from Starting Cohort 6 (Adults). The information (»*respondent's country of birth*«) was surveyed during the interview. The collected information was recoded during data processing to different derived variables. The first of those is variable t405010_g1R, which contains the complete list of countries (lower left, output is shortened to enhance readability). Also, a second variable t405010_g2 was generated, using a more coarse list. This is helpful if some information was collected, but it is too inaccurate to identify the country (e.g., if someone states he was born in North America, you do not know if he meant U.S. or Canada). This second variable is transported to the Download-SUF unaltered (upper right), which means it in some way represents the anonymized version. The first variable t405010_g1R with the full information is purged in the Download-SUF.

RemoteNEPS	Download

```
SC6_pTarget_R_5-1-0.dta
(NEPS SUF, SC6 5.1.0 (remote); doi:10.5157/NEPS:SC6:5.1.0)

t405010_g2 -- Country of birthplace (categorized)
-----------------------------------------------------------
-
                                          |     Freq.
------------------------------------------+----------
-
Valid  -97 Refused                        |         2
       -54 Missing by design              |     36417
       -20 Foreign, but not codeable      |         5
        1  Germany                         |     15388
        2  Italy                           |        52
        3  Poland                          |       244
        4  Romania                         |        98
        5  Turkey                          |       182
        6  Former Yugoslavia               |       113
        7  Former Soviet Union             |       522
        8  Central and South America, Caribbean |   42
        9  Northern and Western Europe     |       154
       10  North America                   |        33
       12  Other Middle East and North Africa |    73
       13  Other Africa                    |        30
       14  Other Asia                      |        74
       15  Other Central and Eastern Europe |       84
       16  Other Southern Europe           |        44
           Total                           |     53557
-----------------------------------------------------------
-

t405010_g1R -- Country of place of birth
-----------------------------------------------------------
                                          |     Freq.
------------------------------------------+----------
Valid  -97 Refused                        |         2
       -54 Missing by design              |     36417
       -20 Foreign, but not codeable      |         5
        0  germany                         |     15388
      120  yugoslavia                      |        30
      121  albania                         |         4
      122  bosnia and herzegovina          |        22
      124  belgium                         |         5
        :                                  |         :
      477  uzbekistan                      |         6
      479  china                           |         7
      482  malaysia                        |         2
      499  other asia                      |         1
           Total                           |     53557
-----------------------------------------------------------
```

```
SC6_pTarget_D_5-1-0.dta
(NEPS SUF, SC6 5.1.0 (download); doi:10.5157/NEPS:SC6:5.1.0)

t405010_g2 -- Country of birthplace (categorized)
-----------------------------------------------------------
-
                                          |     Freq.
------------------------------------------+----------
-
Valid  -97 Refused                        |         2
       -54 Missing by design              |     36417
       -20 Foreign, but not codeable      |         5
        1  Germany                         |     15388
        2  Italy                           |        52
        3  Poland                          |       244
        4  Romania                         |        98
        5  Turkey                          |       182
        6  Former Yugoslavia               |       113
        7  Former Soviet Union             |       522
        8  Central and South America, Caribbean |   42
        9  Northern and Western Europe     |       154
       10  North America                   |        33
       12  Other Middle East and North Africa |    73
       13  Other Africa                    |        30
       14  Other Asia                      |        74
       15  Other Central and Eastern Europe |       84
       16  Other Southern Europe           |        44
           Total                           |     53557
-----------------------------------------------------------
-

t405010_g1R -- Country of place of birth
-----------------------------------------
                                |     Freq.
--------------------------------+----------
Valid  -54 Missing by design    |     36417
       -53 Anonymized           |     17140
           Total                |     53557
-----------------------------------------
```

3.3 Countries (binary variable)

In school cohorts, an aggregation of countries is coarsened to the binary information German/ Non-German. To illustrate this, have a look at variables from Starting Cohort 2 (Kindergarten). The information (»*country of birth target child*«) was surveyed during the interview using a two-step procedure: first, the respondent is asked if the child was born in Germany or abroad (variable p406000, top left). This variable is not subject to anonymization and therefore does not have a suffix and is equally existent in both SUF-versions. If the child was born abroad, a second question regarding the specific country followed (variable p406010_g1R, lower left, output is reduced to enhance readability). This second question is purged in the Download-SUF (lower right). There is no need to generate an additional anonymized version of this second question, as the information is already present in the first shown variable (p406000).

RemoteNEPS	Download

```
SC2_pParent_R_3-0-0.dta
(NEPS SUF, SC2 3.0.0 (remote); doi:10.5157/NEPS:SC2:3.0.0)

p406000 -- Target child born in Germany?
------------------------------------
                        |     Freq.
------------------------+-----------
Valid  -98 Don't know   |         1
        1   Yes         |      7658
        2   No          |       173
        Total           |      7832
Missing .               |      3545
Total                   |     11377
------------------------------------


p406010_g1R -- Country of birth target child
--------------------------------------------------------
                                        |     Freq.
----------------------------------------+-----------
Valid  -98 Don't know                   |         1
       -20 Abroad, but not codeable     |         1
        0    germany                    |      7658
       124 belgium                      |         2
       125 bulgaria                     |         5
        :                               |         :
       475 syrian arab republic         |         2
       476 thailand                     |         2
       479 china                        |         4
       536 new zealand                  |         1
       Total                            |      7832
Missing .                               |      3545
Total                                   |     11377
--------------------------------------------------------
```

```
SC2_pParent_D_3-0-0.dta
(NEPS SUF, SC2 3.0.0 (download); doi:10.5157/NEPS:SC2:3.0.0)

p406000 -- Target child born in Germany?
------------------------------------
                        |     Freq.
------------------------+-----------
Valid  -98 Don't know   |         1
        1   Yes         |      7658
        2   No          |       173
        Total           |      7832
Missing .               |      3545
Total                   |     11377
------------------------------------


p406010_g1R -- Country of birth target child
------------------------------------
                        |     Freq.
------------------------+-----------
Valid  -53 Anonymized   |      7832
Missing .               |      3545
Total                   |     11377
------------------------------------
```

3.4 Federal states

Regional locations are usually recoded to different derived variables (see section 2.3). In this example, we focus on the region of the school in Starting Cohort 3 (5th graders). The first derived variable is a binary variable for Western and Eastern Germany (tx80109_g1), present in both SUF-versions (upper left and upper right). The second variable tx80109_g2R, containing the federal states, is only complete in the RemoteNEPS-SUF (lower left), and is purged in the Download-SUF (lower right).

RemoteNEPS	Download

```
SC3_CohortProfile_R_3-1-0.dta
(NEPS SUF, SC3 3.1.0 (remote); doi:10.5157/NEPS:SC3:3.1.0)

tx80109_g1 -- Sample: Federal state school (west/east)
--------------------------------------------------
                                     |    Freq.
-------------------------------------+------------
Valid   -55 Not determinable         |     458
        1   Western Germany          |   16299
        2   Eastern Germany incl. Berlin |  3784
        Total                        |   20541
--------------------------------------------------

tx80109_g2R -- Sample: Federal state school
--------------------------------------------------
                                     |    Freq.
-------------------------------------+------------
Valid   -55 Not determinable         |     458
        1   Schleswig-Holstein       |     784
        2   Hamburg                  |     377
        3   Lower Saxony             |    2209
        4   Bremen                   |     158
        5   North Rhine-Westphalia   |    4398
        6   Hesse                    |    1308
        7   Rhineland-Palatinate     |     971
        8   Baden-Wuerttemberg       |    3265
        9   Bavaria                  |    2677
        10  Saarland                 |     152
        11  Berlin (Complete)        |     404
        12  Brandenburg              |    1143
        13  Mecklenburg-Western Pomerania |  619
        14  Saxony                   |     617
        15  Saxony-Anhalt            |     401
        16  Thuringia                |     600
        Total                        |   20541
--------------------------------------------------
```

```
SC3_CohortProfile_D_3-1-0.dta
(NEPS SUF, SC3 3.1.0 (download); doi:10.5157/NEPS:SC3:3.1.0)

tx80109_g1 -- Sample: Federal state school (west/east)
--------------------------------------------------
                                     |    Freq.
-------------------------------------+------------
Valid   -55 Not determinable         |     458
        1   Western Germany          |   16299
        2   Eastern Germany incl. Berlin |  3784
        Total                        |   20541
--------------------------------------------------

tx80109_g2R -- Sample: Federal state school
--------------------------------------------------
                                     |    Freq.
-------------------------------------+------------
Valid   -53 Anonymized               |   20541
--------------------------------------------------
```

3 Examples

3.5 Class size

Sometimes, variables are collected numerical. In the example below, the number of students in class was surveyed from the educator in Starting Cohort 2 (Kindergarten) without any further regulations. Thus, the resulting variable e227400_g1R (lower left) does not have a coding scheme. This variable is aggregated to the variable e227400_g1D containing six different class-sizes (upper left). You find this aggregated version in both SUF-versions, the full numerical variable is purged in the Download-SUF, though (lower right).

RemoteNEPS	Download

```
SC2_pEducator_R_3-0-0.dta
(NEPS SUF, SC2 3.0.0 (remote); doi:10.5157/NEPS:SC2:3.0.0)

e227400_g1D -- Class: number of students, total (coarsed)
-----------------------------------------
                                | Freq.
--------------------------------+--------
Valid  -90 Unspecific missing |     10
       -54 Missing by design  |   1803
         1  Below 10          |      3
         2  10 to 14          |     26
         3  15 to 19          |    203
         4  20 to 24          |    450
         5  25 to 29          |    169
         6  30 to 34          |      4
            Total             |   2668
-----------------------------------------

e227400_g1R -- Class: number of students total
-----------------------------------------
                                | Freq.
--------------------------------+--------
Valid  -90 Unspecific missing |     10
       -54 Missing by design  |   1803
         8                    |      3
        10                    |      1
        11                    |      1
        12                    |      4
        13                    |      8
        14                    |     12
        15                    |     21
        16                    |     22
        17                    |     34
        18                    |     58
        19                    |     68
        20                    |     75
        21                    |     89
        22                    |     98
        23                    |     88
        24                    |    100
        25                    |     83
        26                    |     39
        27                    |     27
        28                    |     14
        29                    |      6
        30                    |      3
        31                    |      1
            Total             |   2668
-----------------------------------------
```

```
SC2_pEducator_D_3-0-0.dta
(NEPS SUF, SC2 3.0.0 (download); doi:10.5157/NEPS:SC2:3.0.0)

e227400_g1D -- Class: number of students, total (coarsed)
-----------------------------------------
                                | Freq.
--------------------------------+--------
Valid  -90 Unspecific missing |     10
       -54 Missing by design  |   1803
         1  Below 10          |      3
         2  10 to 14          |     26
         3  15 to 19          |    203
         4  20 to 24          |    450
         5  25 to 29          |    169
         6  30 to 34          |      4
            Total             |   2668
-----------------------------------------

e227400_g1R -- Class: number of students total
-----------------------------------------
                                | Freq.
--------------------------------+--------
Valid  -54 Missing by design  |   1803
       -53 Anonymized         |    865
            Total             |   2668
-----------------------------------------
```

A Appendix

A.1 Complete overview

The following table lists all variables which are affected by anonymization methods. For all variables, full information is available OnSite, although for many of them, you already find this complete information in the RemoteNEPS-version. Only for those variables marked with an asterisk, the full information is only present OnSite. For most of the variables, an aggregated or generated version is available, which still transports some information to the Download-version. Please see the previous chapters for more information.

Table 7: Complete overview of all restricted variables

file	variable	label (*only available OnSite)
pParent	p32601a	Position generator: country nurse or male nurse
pParent	p32601b	Position generator: country engineer
pParent	p32601c	Position generator: country warehouse/transport worker
pParent	p32601d	Position generator: country social worker
pParent	p32601e	Position generator: country sales clerk
pParent	p32601f	Position generator: country police officer
pParent	p32601g	Position generator: country doctor
pParent	p32601h	Position generator: country banker
pParent	p32601k	Position generator: country car mechanic
pParent	p32601l	Position generator: country legal practitioner
pParent	p32601m	Position generator: country optician
pParent	p32601n	Position generator: country translator
pParent	p32601o	Position generator: country teacher at elementary school, Hauptschule or Real...
pParent	p400010_g1	Country of birth respondent
pParent	p400010_g2	Country of birth respondent (aggregated)
pParent	p400070_g1	Country of birth mother respondent
pParent	p400070_g2	Country of birth respondent's mother (aggregated)
pParent	p400090_g1	Respondent's father's country of birth
pParent	p400090_g2	Country of birth respondent's father (aggregated)
pParent	p400500_g3	group of origin
pParent	p400500_g3v1	Group of origin
pParent	p401150_g2	Nationality Respondent not German (aggregated)
pParent	p403010_g1	Country of birth partner abroad
pParent	p403010_g2	Country of birth partner (aggregated)
pParent	p403070_g1	Country of birth mother partner
pParent	p403070_g2	Country of birth partner's mother (aggregated)
pParent	p403090_g1	Country of birth father partner
pParent	p403090_g2	Country of birth partner's father (aggregated)
pParent	p404050_g2	Other nationality partner (aggregated)
pParent	p406010_g1	Country of birth target child
pParent	p406010_g2	Country of birth target child (aggregated)
pParent	p407050_g2	Nationality Target child (aggregated)

(...)

Table 7: (continued)

file	variable	label (*only available OnSite)
pParent	p407060_g2	Target child's second nationality (aggregated)
pParent	p731804_g1	Highest educational qualification of respondent abroad (country)
pParent	p731804_g2	Country of respondent's highest educational qualification (aggregated)
pParent	p731823_g1	Country of vocational qualification (additional response)
pParent	p731823_g2	Country of professional qualification (additional response, aggregated)
pParent	p731854_g1	Country of partner's highest school-leaving qualification
pParent	p731854_g2	Country of highest educational qualification partner (aggregated)
pParent	p731873_g1	Country of vocational qualification, partner, (additional response)
pParent	p731873_g2	Country of partner's professional qualification (add. response, aggregated)
pTarget	t400000_g1	Birth country
pTarget	t400000_g2	Country of birth (aggregated)
spParentSchool	p723060_g1	Country of school
spParentSchool	p723060_g2	Country of school (aggregated)
ParentMethods	px80204	Interview: Interview language (realized case)
TargetMethods	tx80204	Interview: Interview language (realized case)
pEducator	e41100a_g2	Mother tongue (response 1, ISO 639.2)
pEducator	e41100a_g3	Mother tongue (response 2, ISO 639.2)
pEducator	e41100a_g4	Mother tongue (response 3, ISO 639.2)
pEducator	e41100a_g5	Mother tongue (response 4, ISO 639.2)
pParent	p410000_g1	First language/mother tongue child (ISO 639.2)
pParent	p410000_g2	First language/mother tongue child (aggregated)
pParent	p410002_g1	Further first language/ mother tongue child (ISO 639.2)
pParent	p410002_g2	Further first language/ mother tongue child (aggregated)
pParent	p412001_g1	Interaction language household, detailed (ISO 639.2)
pParent	p412001_g2	Interactive language household detailed (aggregated)
pParent	p413000_g1	First language/mother tongue respondent (ISO 639.2)
pParent	p413000_g2	First language/mother tongue respondent (aggregated)
pParent	p413002_g1	Further first language/ mother tongue respondent (ISO 639.2)
pParent	p413002_g2	Further first language/ mother tongue respondent (aggregated)
pParent	p414000_g1	First language/mother tongue partner (ISO 639.2)
pParent	p414000_g2	First language/mother tongue partner (aggregated)
pParent	p414002_g1	Further first language/ mother tongue partner (ISO 639.2)
pParent	p414002_g2	Further first language/ mother tongue partner (aggregated)
pParent	p414100_g1	First language/mother tongue new partner (ISO 639.2)
pParent	p414100_g2	First language/mother tongue new partner (aggregated)
pParent	p414102_g1	Further first language/ mother tongue new partner (ISO 639.2)
pParent	p414102_g2	Further first language/ mother tongue new partner (aggregated)
pTarget	t41000a_g2	Mother tongue (response 1, ISO 639.2)

(...)

Table 7: (continued)

file	variable	label (*only available OnSite)	
pTarget	t41000a_g3	Mother tongue (response 2, ISO 639.2)	
pTarget	t41000a_g4	Mother tongue (response 3, ISO 639.2)	
pTarget	t41000a_g5	Mother tongue (response 4, ISO 639.2)	
pParent	p401150_g1	Citizenship respondent not German	
pParent	p404050_g1	Other citizenship partner	
pParent	p407050_g1	Citizenship of the target child	
pParent	p407060_g1	Second citizenship of the target child	
pCourseClass	ec20020	My lessons have no connection to a course, I use an open concept, specifically:	*
pEducator	e212819	Further training, other, text	*
pEducator	e212823	Group leader: further education, additional certification, text	*
pEducator	e21980b	Kindergarten grp leaqder:professional qualification,other qualification,text	*
pEducator	e219817	Group leader: extent of work - other, text	*
pEducator	ec03091	Working hours: Other text	*
pGroups	e217406	Other developmental disorders (open)	*
pInstitution	h212013	Kindergarten: participation in quality improvement measure, other measure, text)	*
pInstitution	h212307	Kindergarten: frequency of honorary tasks, other, text	*
pInstitution	h212819	Training, other, text	*
pInstitution	h217006	Kindergarten: number of children with other diagnosed developm. disorders, text	*
pInstitution	h218020	Kindergarten: collaboration with a school: other, text	*
pInstitution	h21905t	Kindergarten: problems with surroundings, other, text	*
pInstitution	h21920a	Kindergarten: diagnostic offers, other, text	*
pInstitution	h21920b	Kindergarten: therapeutic offers, other, text	*
pInstitution	h21920c	Kindergarten: offers for parents, other, text	*
pInstitution	h21932t	Other qualification, text	*
pInstitution	h21980a	Kindergarten management: professional qualification: other qualification: Text	*
pInstitution	h219817	Use of actual weekly working hours, other, text	*
pInstitution	h22802t	School: cooperation with kindergarten: other measures, text	*
pInstitution	h22803t	School: Cooperation with secondary schools: Other – text	*
pInstitution	h229003	School: approach, educational, text	*
pInstitution	h229005	School: approach, promotion, text	*
pInstitution	h229007	School: approach, integration, text	*
pInstitution	h229009	School: approach, other, text	*
pInstitution	h22903t	School: problems in the environment: other, text	*
pInstitution	h40184a	Measures for parents with an immigration background open 1	*
pInstitution	h40185a	Measures intercultural competence open for educators (open)	*
pInstitution	h416110	Remedial language teaching - open	*
pInstitution	h417120	Migration-specific support measures - open	*
pInstitution	h417160	Migration-specific additional support for teachers: Open specification	*

(...)

Table 7: (continued)

file	variable	label (*only available OnSite)	
pInstitution	h418012	Language remediation: individual remediation - other, open	*
pInstitution	h418052	Language remediation: small group remediation - other, open	*
pInstitution	h418092	Language remediation: whole group remediation - other, open	*
pInstitution	hb1006t	Execution of the speech remediation persons with other qualification, text	*
pInstitution	hb1015e	Other from of bilingual care	*
pInstitution	hc1007d	Remedial language teaching: other type of assistance (open)	*
pInstitution	hd0041b	No entry made	*
pParent	p19020j	Special needs OPEN	*
pParent	p262605	Outside activities: Usage other, Text	*
pParent	p401001	Migrant status Other	*
pParent	p403041	Migrant status partner - other	*
pParent	p524853	AD(H)D: Medications - open	*
pParent	p723090	Type of school (open)	*
pParent	p731803	Highest educational qualification Respondent, type open	*
pParent	p731814	Vocational qualification respondent (open)	*
pParent	p731817	Type tertiary qualification respondent (open)	*
pParent	p731853	Highest educational qualification Partner, type open	*
pParent	p731864	Professional qualification partner (open)	*
pParent	pc13140	Deferral of enrollment: attendance at current facility, other	*
pTarget	t27111d	Student: Courses outside of school: Other courses: Text, where	*
pTarget	t400071	Other country of birth of the mother	*
pTarget	t400091	Other country of birth of the father	*
pTarget	t410001	Language of origin of TP - open	*
pTarget	t435030	other religion or denomination	*
pTarget	t435050	Religion or denomination (other)	*
spParentGap	ps29102	Other activity	*
spParentSchool	p723090	Type of school (open)	*
spSibling	p723090	Type of school (open)	*
CohortProfile	tx80109_g2	Sample: Federal state school	
WeightsElementarySchool	stratum_imp2	implicit stratum 2 (federal states of institution)	
WeightsElementarySchool	stratum_imp3	implicit stratum 3 (regional classification of institution)	
WeightsElementarySchool	stratum_imp4	implicit stratum 4 (organizing institution)	
pParent	p723050	Federal state of the school	
pParent	p751001_g2	Place of Residence (Federal State)	
pParent	pb11610_g2	Place of elementary school (federal state)	
spParentSchool	p723030_g2	Municipality of school (Federal state)	
spSibling	p723050	State of school	
pCourseClass	e227400_g1	Class: number of students total	
pCourseClass	e229400	Class: Facilities: Classroom size	
pCourseClass	ec0150a	Number of lessons hours per week: German/literacy	
pCourseClass	ec0151a	Number of lessons per week: Math	

(...)

Table 7: (continued)

file	variable	label (*only available OnSite)
pCourseClass	ec0152a	Number of lessons per week: General knowledge/local history and geography
pCourseClass	ec01530	Number of lessons/week in total for the class
pEducator	e219820	Group leader: duration of professional activity: previous facilities
pEducator	e219821	Group leader: duration of professional activity: current facility
pEducator	e76212y	Year of birth
pGroups	e219410	Kindergarten core group: room use: number of rooms
pGroups	e219411	Kindergarten core group: room use: size of rooms
spChildCare	fee	Care: fees (euros)
CohortProfile	tx80505	List of children / students: Special educational needs Target child
pEducator	e219800	Kindergarten group leader: professional qualification
pParent	p731911	Number of employees respondent
pParent	p731961	Number of employees partner
WeightsElementarySchool	tx80112	Amount of students in grade 2 total (accord. to the Federal Statistical Office)
pCourseClass	e19001c	Number students with special needs
pCourseClass	e199001	Number special needs classes: learning
pCourseClass	e199002	Number special needs classes: language
pCourseClass	e199003	Number special needs classes: physical/motor development
pCourseClass	e199004	Number special needs classes: emotional/social development
pCourseClass	e199005	Number special needs classes: mental development
pCourseClass	e199006	Number special needs classes: seeing
pCourseClass	e199007	Number special needs classes: hearing
pCourseClass	e199008	Number special needs classes: autism
pCourseClass	e227400	Class: number of female students
pCourseClass	e227401	Class: number of male students
pCourseClass	e451000	Daycare: number of students with migrant background
pCourseClass	e79201a	Class: share class affiliation - low class
pCourseClass	e79201b	Class: share class affiliation - middle class
pCourseClass	e79201c	Class: share class affiliation - upper class
pCourseClass	e79202a	Class: share parents with a higher education degree
pEducator	e536020	Age Career choice
pEducator	e53702y	Year of university entrance qualification
pEducator	e537030	Federal state higher education entrance qualification (Federal state)
pEducator	e53710y	Start Teaching degree course
pEducator	e537150	Year of state examination
pEducator	e537200	Federal state second Staatsexamen (Federal state)
pEducator	e76212m	Month of birth *
pGroups	e217401	Kindergarten core group: girls registered
pGroups	e217402	Kindergarten core group: boys registered
pGroups	e217403	Core group: children with diagnosed speech disorders (number)

(...)

Table 7: (continued)

file	variable	label (*only available OnSite)
pGroups	e217404	Core group: children with diagnosed behavioral disorders (number)
pGroups	e217405	Children with other developmental disorders (number)
pGroups	e217411	Core group: year of birth 2010 or later; number of children altogether
pGroups	e217412_w1	Core group: year of birth 2009 and later; number of children altogether
pGroups	e217412_w2	Core group: year of birth 2009; number of children, altogether
pGroups	e217413	Core group: year of birth 2008; number of children, altogether
pGroups	e217414	Core group: year of birth 2007; number of children altogether
pGroups	e217415	Core group: year of birth 2006; number of children altogether
pGroups	e217416_w1	Core group: year of birth 2005; number of children, altogether
pGroups	e217416_w2	Core group: year of birth 2005 or earlier: number of children altogether
pGroups	e217417	Core group: year of birth 2004 and earlier; number of children, altogether
pGroups	e217421	Core group: year of birth 2010 or later; number children attending up to 5 hours
pGroups	e217422_w1	Core group: year of birth 2009 and later; number children, attend up to 5 hours
pGroups	e217422_w2	Core group: year of birth 2009; number of children attending up to 5 hours
pGroups	e217423	Core group: year of birth 2008; number of children, attending up to 5 hours
pGroups	e217424	Core group: year of birth 2007; number of children, attending up to 5 hours
pGroups	e217425	Core group: year of birth 2006; number of children attending up to 5 hours
pGroups	e217426_w1	Core group: year of birth 2005; number of children, attend up to 5 hours
pGroups	e217426_w2	Core group: year of birth 2005 or earlier: number children, attending up to 5h
pGroups	e217427	Core group: year of birth 2004 and earlier; number of children, attend up to 5 h
pGroups	e217431	Core group: year of birth 2010 or later; number of children attending 5-7 hours
pGroups	e217432_w1	Core group: year of birth 2009 and later; number of children, attend 5-7 hours
pGroups	e217432_w2	Core group: year of birth 2009; number of children attending 5-7 hours
pGroups	e217433	Core group: year of birth 2008; number of children, attending 5-7 hours
pGroups	e217434	Core group: year of birth 2007; number of children, attending 5-7 hours

(...)

Table 7: (continued)

file	variable	label (*only available OnSite)
pGroups	e217435	Core group: year of birth 2006; number of children, attending 5-7 hours
pGroups	e217436_w1	Core group: year of birth 2005; number of children, attend 5-7 hours
pGroups	e217436_w2	Core group: year of birth 2005 or earlier; number children, attending 5-7 hours
pGroups	e217437	Core group: year of birth 2004 and earlier; number of children, attend 5-7 hours
pGroups	e217441	Core group: year of birth 2010 or later; number children attending more than 7h
pGroups	e217442_w1	Core group: year of birth 2009 and later; nr. children, attend more than 7 hours
pGroups	e217442_w2	Core group: year of birth 2009; number of children attending more than 7 hours
pGroups	e217443	Core group: year of birth 2008; number of children, attending more than 7 hours
pGroups	e217444	Core group: year of birth 2007; number of children, attending more than 7 hours
pGroups	e217445	Core group: year of birth 2006; number of children, attending more than 7 hours
pGroups	e217446_w1	Core group: year of birth 2005; number of children, attend more than 7 hours
pGroups	e217446_w2	Core group: year of birth 2005 or earlier; nr. children, attending more than 7h
pGroups	e217447	Core group: Year of birth 2004 and earlier; nr. children, attend over 7 hours
pGroups	e217451	Core group: year of birth 2010 and later; number of children, disability
pGroups	e217452_w1	Core group: year of birth 2009 and later; number of children, disability
pGroups	e217452_w2	Core group: year of birth 2009; number of children, disability
pGroups	e217453	Core group: year of birth 2008; number of children, disability
pGroups	e217454	Core group: year of birth 2007; number of children, disability
pGroups	e217455	Core group: year of birth 2006; number of children, disability
pGroups	e217456_w1	Core group: year of birth 2005; number of children, disability
pGroups	e217456_w2	Core group: year of birth 2005 or earlier; number of children, disability
pGroups	e217457	Core group: year of birth 2004 and earlier; number of children, disability
pGroups	e401200	Number of children with another interactive language
pGroups	e45110a	Core group: year of birth 2004 and earlier; number of children, mig. background
pGroups	e45110b_w1	Core group: year of birth 2005; number of children, migration background

(...)

Table 7: (continued)

file	variable	label (*only available OnSite)	
pGroups	e45110b_w2	Core group: year of birth 2005 or earlier; number children, migrant background	
pGroups	e45110c	Core group: year of birth 2006; number of children, migrant background	
pGroups	e45110d	Core group: year of birth 2007; number of children, migrant background	
pGroups	e45110e	Core group: year of birth 2008; number of children, migrant background	
pGroups	e45110f_w1	Core group: year of birth 2009 and later; number children, migration background	
pGroups	e45110f_w2	Core group: year of birth 2009; number of children with migrant background	
pGroups	e45110g	Core group: year of birth 2010 or later; number children with migrant background	
pGroups	e79101a	Group: proportion class affiliation - low class	
pGroups	e79101b	Group: proportion class affiliation - middle class	
pGroups	e79101c	Group: proportion class affiliation - upper class	
pGroups	e79102a	Group: proportion university graduates	
pGroups	eb1000a	Language skills above average age - all children	
pGroups	eb1000b	Language skills above average age-thereof children with migration background*	
pGroups	eb1000c	Language skills average - all children	
pGroups	eb1000d	Language skills average - portion of children with migrant background*	
pGroups	eb1000e	Language skills below average - all children	
pGroups	eb1000f	Language skills below average - portion of children with migrant background*	
pInstitution	h76512m	Month of birth	*
pParent	p435030_g1	Religion and religiousness: Other religious community	
pTarget	t400070	Country of birth of the mother	
pTarget	t400090	Country of birth of the father	
pTarget	t435030_g1	Religion and religiousness: other religious community	
pTarget	t723400	grade attended	*
pTargetMicrom			*
pTargetRegioInfas			*
pInstitutionMicrom			*
pInstitutionRegioInfas			*
pInstitution			

B References