

Starting Cohort 1: Newborns (SC1)
SUF Version 3.0.0
Anonymization Procedures
Tobias Koberg

Copyrighted Material
Leibniz Institute for Educational Trajectories (LifBi)
Wilhelmsplatz 3, 96047 Bamberg
Director: Prof. Dr. Hans-Günther Roßbach
Executive Director of Research: Dr. Jutta von Maurice
Executive Director of Administration: Dr. Robert Polgar
Bamberg, 2016

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) SC1, version 3-0-0 (doi:10.5157/NEPS:SC1:3.0.0). 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 > SC1 > 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 Lane, Heus, and Mulcahy, 2008, 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 oncoming publication Koberg, 2016. 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

1 Preamble

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 mode¹, 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

2 Conducted measures

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.

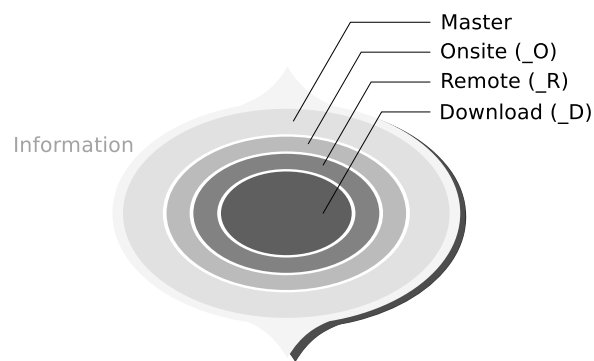


Figure 1: Onion-shaped model defining the different anonymization levels

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 99 out of 2150 variables are modified in some way.

2 Conducted measures

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
MethodsCAPI	px80218	Interview: other reason for an interview with father (open)
MethodsDirectMeasures	px01007	Video recording did not work for 'Looking at pictures' cause open
MethodsDirectMeasures	px02007	Video recording did not work for 'Playing with mother' cause open
MethodsDirectMeasures	px03005	performance did not work for 'Playing with the female interviewer': cause open
MethodsDirectMeasures	px04016	Other reasons for refusing the video recording
MethodsDirectMeasures	px04028	Game not performed: other reasons
MethodsDirectMeasures	tx80637	Habituation: task 1 - comments
pEducator	e20110t	Day care facility: Frequency: Other things [open]
pEducator	e209103	Youth/child care worker: Vocational degree: Other text

(...)

2 Conducted measures

Table 2: (continued)

file	variable	label
pEducator	e209125	Day care institution: Amount of working hours: Other, text
pEducator	e209150	Day care institution: Opening hours: Opening time
pEducator	e209151	Day care institution: Opening hours: Closing time
pEducator	e209160	Day care institution: Opening hours: Opening time
pEducator	e209161	Day care institution: Opening hours: Closing time
pEducator	e209170	Day care institution: Opening hours: Opening time
pEducator	e209171	Day care institution: Opening hours: Closing time
pEducator	e209180	Day care institution: Opening hours: Opening time
pEducator	e209181	Day care institution: Opening hours: Closing time
pEducator	e53902t	Day care facility: Admission criteria: Other criteria text
pEducator	ea01040	Day care facility: Acclimatization period: Other
pEducator	ea0104b	Day care institution: Adjustment period: Presence of parents, other
pEducator	ea0201g	Day care facility: Conversations with parents: Other issues, including
pEducator	ea0301b	Day care facility: Provider text
pEducator	ea0401f	Day care facility: Fees: Other text
pEducator	ea0502b	Day care institution: Speech promotion: Other, text
pEducator	ea0505e	Day care facility: Conduct of support: Other qualifications
pEducator	ea0715c	Day care institution: Group leader: Advanced training, other, text
pEducator	ea0801b	Day care institution: Further education, professional qualification, text
pEducator	ea0802b	Day care institution: Further education, certified additional qualification, ...
pEducator	ea1301g	Day care facility: Financial source: Other text
pEducatorChildminder	e20110t	Tageseinrichtung: Frequency of activity: Other [open]
pEducatorChildminder	e209118	Day care: Professional qualification, other qualification, text
pEducatorChildminder	e20912t	Day care: extent working hours: other, text
pEducatorChildminder	e209144	Day care: Duration occupation other, text
pEducatorChildminder	e20915a	Day care: opening hours: opening time
pEducatorChildminder	e20915b	Day care: opening hours: closing time
pEducatorChildminder	e20916a	Day care: opening hours: opening time
pEducatorChildminder	e20916b	Day care: opening hours: closing time
pEducatorChildminder	e20917a	Day care: opening hours: opening time
pEducatorChildminder	e20917b	Day care: opening hours: closing time
pEducatorChildminder	e20918a	Day care: opening hours: opening time
pEducatorChildminder	e20918b	Day care: opening hours: closing time
pEducatorChildminder	e53902t	Day care: Admission criteria: Other criteria text
pEducatorChildminder	e540000t	Day care: Employment relationship: Type text
pEducatorChildminder	ea2104b	Day care: adjustment period: presence of parents, other
pEducatorChildminder	ea2201g	Day care: Discuss with parents: Other topics, text
pEducatorChildminder	ea2401f	Day care: Fees: Other, text
pEducatorChildminder	ea2715c	Day care: further training, other, text

(...)

2 Conducted measures

Table 2: (continued)

file	variable	label
pParent	p271603	Parent child groups offer: Open 1
pParent	p271604	Parent child groups offer: Open 2
pParent	p271605	Parent child groups offer: Open 3
pParent	p271606	Parent child groups offer: Open 4
pParent	p271607	Parent child groups offer: Open 5
pParent	p401001	Other reasons behind relocation to Germany
pParent	p403041	Other reasons behind partner's relocation to Germany
pParent	p435030	Religion and religiosity: other religious community
pParent	p524410	Nature of health impairment: Other
pParent	p524428	Type of health impairment: other
pParent	pa10002	ELFRA2: syntax item 2
pParent	pa10048	ELFRA2: syntax item 48

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 into larger categories – mainly allocating continents and, in detail, states in Europe. For a complete recoding map of all countries/languages/nationalities, see appendix section A.2. This recoding rule was initially based on the methods used for the IAB ALWA study (Antoni et al., 2010) and the recommendations given by Müller, Blien, and Knoche, 1991. It has been updated and adjusted to better serve the data user's needs. Please note that this aggregated version is generated not during the anonymization process, but as an additional derived variable. Therefore, you do not find a variable suffixed `_D`, but two derived variables (suffixed `_g1` and `_g2`), where only the first has an anonymization suffix (i.e. `_g1R`), indicating that the full information is available in RemoteNEPS only.

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

file	variable	label
pParent	p400010_g1	Country of birth respondent
pParent	p400070_g1	Respondent's mother's country of birth
pParent	p400090_g1	Respondent's father's country of birth
pParent	p403010_g1	Country of birth partner abroad
pParent	p403070_g1	Partner's mother's country of birth
pParent	p403090_g1	Partner's father's country of birth
pParent	p406010_g1	Country of birth target child
pParent	p731804_g1	Highest educational qualification of respondent abroad (country)

(...)

2 Conducted measures

Table 3: (continued)

file	variable	label
pParent	p731854_g1	Highest educational qualification partner abroad (country)
pParent	p412001_g1	Interactive language household detailed (ISO 639.2)
pParent	p412071_g1	Interactive language interviewed parent - child (ISO 639.2)
pParent	p412081_g1	Interaction language partner - child (ISO 639.2)
pParent	p413000_g1	Respondent's language of origin (ISO 639.2)
pParent	p413002_g1	Respondent's other language of origin (ISO 639.2)
pParent	p414000_g1	Partner's language of origin (ISO 639.2)
pParent	p414002_g1	Partner's other language of origin (ISO 639.2)
pParent	p401150_g1	Citizenship respondent not German
pParent	p404050_g1	Other citizenship partner
pParent	p407050_g1	Citizenship - target child
pParent	p407060_g1	Target child's second nationality

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
Weights	px80101	Sample: Federal State
pParent	p751001_g2	Place of Residence (RS Federal State)

2 Conducted measures

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) `pEducator` `pEducatorChildminder`, 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

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.

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

[file] variable: label

RemoteNEPS	Download
	1 under 2000g
	2 2000g to under 2500g
	3 2500g to under 3000g
	4 3000g to under 3500g
	5 3500g to under 4000g
	6 4000g to under 4500g
	7 4500g and more

(...)

2 Conducted measures

Table 5: (continued)

[file] variable: label

[pParent] p529001: height at birth (aggregated)

RemoteNEPS		Download
	1	under 44cm
	2	44cm to under 46cm
	3	46cm to under 48cm
	4	48cm to under 50cm
	5	50cm to under 52cm
	6	52cm to under 54cm
	7	54cm to under 56cm
	8	56cm and more

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

RemoteNEPS		Download
1 to 4	1	1 to 4
5 to 9	2	5 to 9
10 to 19	3	10 to 19
20 to 49	4	20 and more
50 to 99	5	
100 to 199	6	
200 to 249	7	
250 to 499	8	
500 to 999	9	
1,000 to 1,999	10	
2,000 and more	11	

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

RemoteNEPS		Download
1 to 4	1	1 to 4
5 to 9	2	5 to 9
10 to 19	3	10 to 19
20 to 49	4	20 and more
50 to 99	5	
100 to 199	6	
200 to 249	7	
250 to 499	8	
500 to 999	9	
1,000 to 1,999	10	
2,000 and more	11	

[pParent] pa01150: Set-up fee: Euro (aggregated)

RemoteNEPS		Download
	1	Under 100 Euro
	2	100 to under 200 Euro
	3	200 to under 300 Euro
	4	300 to under 400 Euro
	5	400 to under 500 Euro
	6	500 Euro and more

(...)

3 Examples

Table 5: (continued)

[file] variable: label

[pParent] pa01250: Day-care fees: Euro (aggregated)	RemoteNEPS	Download
		1 Under 100 Euro
		2 100 to under 200 Euro
		3 200 to under 300 Euro
		4 300 to under 400 Euro
		5 400 to under 500 Euro
		6 500 Euro and more

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

file	variable	label
pEducator	e209105	Day care institution: Duration of professional activity: Current institution
pEducator	e209111	Day care institution Core group: Room use: Size of rooms
pParent	p435030_g1	Religion and religiosity: other religious community
pParent	pa01151	Facility fees: Cents
pParent	pa01251	Childminder fees: Cents

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.

3.1 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

3 Examples

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)		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)		t731406_D -- Number of mother's employees (categorized)	
	Freq.		Freq.
Valid		Valid	
-98 Don't know	7	-98 Don't know	7
-97 Refused	1	-97 Refused	1
-54 Missing by design	36700	-54 Missing by design	36700
0 none	423	0 none	423
1 1 to 4	330	1 1 to 4	330
2 5 to 9	64	2 5 to 9	64
3 10 to 19	22	3 10 to 19	22
4 20 and more	28	4 20 and more	28
Total	37575	Total	37575
Missing .	15982	Missing .	15982
Total	53557	Total	53557
t731406_R -- Number of mother's employees		t731406_R -- Number of mother's employees	
	Freq.		Freq.
Valid		Valid	
-98 Don't know	7	-54 Missing by design	36700
-97 Refused	1	-53 Anonymized	875
-54 Missing by design	36700	Total	37575
0 None	423	Missing .	15982
1 1 to 4	330	Total	53557
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		

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

3 Examples

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)		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)		t405010_g2 -- Country of birthplace (categorized)	
	Freq.		Freq.
Valid		Valid	
-97 Refused	2	-97 Refused	2
-54 Missing by design	36417	-54 Missing by design	36417
-20 Foreign, but not codeable	5	-20 Foreign, but not codeable	5
1 Germany	15388	1 Germany	15388
2 Italy	52	2 Italy	52
3 Poland	244	3 Poland	244
4 Romania	98	4 Romania	98
5 Turkey	182	5 Turkey	182
6 Former Yugoslavia	113	6 Former Yugoslavia	113
7 Former Soviet Union	522	7 Former Soviet Union	522
8 Central and South America, Caribbean	42	8 Central and South America, Caribbean	42
9 Northern and Western Europe	154	9 Northern and Western Europe	154
10 North America	33	10 North America	33
12 Other Middle East and North Africa	73	12 Other Middle East and North Africa	73
13 Other Africa	30	13 Other Africa	30
14 Other Asia	74	14 Other Asia	74
15 Other Central and Eastern Europe	84	15 Other Central and Eastern Europe	84
16 Other Southern Europe	44	16 Other Southern Europe	44
Total	53557	Total	53557
t405010_g1R -- Country of place of birth		t405010_g1R -- Country of place of birth	
	Freq.		Freq.
Valid		Valid	
-97 Refused	2	-54 Missing by design	36417
-54 Missing by design	36417	-53 Anonymized	17140
-20 Foreign, but not codeable	5	Total	53557
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		

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,

3 Examples

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
<pre>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 -----</pre>	<pre>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 -----</pre>
<pre>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 -----</pre>	<pre>p406010_g1R -- Country of birth target child ----- Freq. -----+----- Valid -53 Anonymized 7832 Missing . 3545 Total . 11377 -----</pre>

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).

3 Examples

RemoteNEPS		Download	
SC3_CohortProfile_R_3-1-0.dta (NEPS SUF, SC3 3.1.0 (remote); doi:10.5157/NEPS:SC3:3.1.0)		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)		tx80109_g1 -- Sample: Federal state school (west/east)	
	Freq.		Freq.
Valid	-55 Not determinable 458	Valid	-55 Not determinable 458
	1 Western Germany 16299		1 Western Germany 16299
	2 Eastern Germany incl. Berlin 3784		2 Eastern Germany incl. Berlin 3784
	Total 20541		Total 20541
tx80109_g2R -- Sample: Federal state school		tx80109_g2R -- Sample: Federal state school	
	Freq.		Freq.
Valid	-55 Not determinable 458	Valid	-53 Anonymized 20541
	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		

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).

3 Examples

RemoteNEPS

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

Download

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	p400010_g1	Country of birth respondent	
pParent	p400070_g1	Respondent's mother's country of birth	
pParent	p400090_g1	Respondent's father's country of birth	
pParent	p403010_g1	Country of birth partner abroad	
pParent	p403070_g1	Partner's mother's country of birth	
pParent	p403090_g1	Partner's father's country of birth	
pParent	p406010_g1	Country of birth target child	
pParent	p731804_g1	Highest educational qualification of respondent abroad (country)	
pParent	p731854_g1	Highest educational qualification partner abroad (country)	
pParent	p412001_g1	Interactive language household detailed (ISO 639.2)	
pParent	p412071_g1	Interactive language interviewed parent - child (ISO 639.2)	
pParent	p412081_g1	Interaction language partner - child (ISO 639.2)	
pParent	p413000_g1	Respondent's language of origin (ISO 639.2)	
pParent	p413002_g1	Respondent's other language of origin (ISO 639.2)	
pParent	p414000_g1	Partner's language of origin (ISO 639.2)	
pParent	p414002_g1	Partner's other language of origin (ISO 639.2)	
pParent	p401150_g1	Citizenship respondent not German	
pParent	p404050_g1	Other citizenship partner	
pParent	p407050_g1	Citizenship - target child	
pParent	p407060_g1	Target child's second nationality	
MethodsCAPI	px80218	Interview: other reason for an interview with father (open)	*
MethodsDirectMeasures	px01007	Video recording did not work for 'Looking at pictures' cause open	*
MethodsDirectMeasures	px02007	Video recording did not work for 'Playing with mother' cause open	*
MethodsDirectMeasures	px03005	performance did not work for 'Playing with the female interviewer': cause open	*
MethodsDirectMeasures	px04016	Other reasons for refusing the video recording	*
			(...)

Table 7: (continued)

file	variable	label (* only available OnSite)	
MethodsDirectMeasures	px04028	Game not performed: other reasons	*
MethodsDirectMeasures	tx80637	Habituation: task 1 - comments	*
pEducator	e20110t	Day care facility: Frequency: Other things [open]	*
pEducator	e209103	Youth/child care worker: Vocational degree: Other text	*
pEducator	e209125	Day care institution: Amount of working hours: Other, text	*
pEducator	e209150	Day care institution: Opening hours: Opening time	*
pEducator	e209151	Day care institution: Opening hours: Closing time	*
pEducator	e209160	Day care institution: Opening hours: Opening time	*
pEducator	e209161	Day care institution: Opening hours: Closing time	*
pEducator	e209170	Day care institution: Opening hours: Opening time	*
pEducator	e209171	Day care institution: Opening hours: Closing time	*
pEducator	e209180	Day care institution: Opening hours: Opening time	*
pEducator	e209181	Day care institution: Opening hours: Closing time	*
pEducator	e53902t	Day care facility: Admission criteria: Other criteria text	*
pEducator	ea01040	Day care facility: Acclimatization period: Other	*
pEducator	ea0104b	Day care institution: Adjustment period: Presence of parents, other	*
pEducator	ea0201g	Day care facility: Conversations with parents: Other issues, including	*
pEducator	ea0301b	Day care facility: Provider text	*
pEducator	ea0401f	Day care facility: Fees: Other text	*
pEducator	ea0502b	Day care institution: Speech promotion: Other, text	*
pEducator	ea0505e	Day care facility: Conduct of support: Other qualifications	*
pEducator	ea0715c	Day care institution: Group leader: Advanced training, other, text	*
pEducator	ea0801b	Day care institution: Further education, professional qualification, text	*
pEducator	ea0802b	Day care institution: Further education, certified additional qualification, ...	*
pEducator	ea1301g	Day care facility: Financial source: Other text	*
pEducatorChildminder	e20110t	Tageseinrichtung: Frequency of activity: Other [open]	*
pEducatorChildminder	e209118	Day care: Professional qualification, other qualification, text	*
pEducatorChildminder	e20912t	Day care: extent working hours: other, text	*
pEducatorChildminder	e209144	Day care: Duration occupation other, text	*
pEducatorChildminder	e20915a	Day care: opening hours: opening time	*
pEducatorChildminder	e20915b	Day care: opening hours: closing time	*
pEducatorChildminder	e20916a	Day care: opening hours: opening time	*
pEducatorChildminder	e20916b	Day care: opening hours: closing time	*
pEducatorChildminder	e20917a	Day care: opening hours: opening time	*
pEducatorChildminder	e20917b	Day care: opening hours: closing time	*
pEducatorChildminder	e20918a	Day care: opening hours: opening time	*
pEducatorChildminder	e20918b	Day care: opening hours: closing time	*
pEducatorChildminder	e53902t	Day care: Admission criteria: Other criteria text	*
pEducatorChildminder	e540000t	Day care: Employment relationship: Type text	*
pEducatorChildminder	ea2104b	Day care: adjustment period: presence of parents, other	*

(...)

Table 7: (continued)

file	variable	label (* only available OnSite)	
pEducatorChildminder	ea2201g	Day care: Discuss with parents: Other topics, text	*
pEducatorChildminder	ea2401f	Day care: Fees: Other, text	*
pEducatorChildminder	ea2715c	Day care: further training, other, text	*
pParent	p271603	Parent child groups offer: Open 1	*
pParent	p271604	Parent child groups offer: Open 2	*
pParent	p271605	Parent child groups offer: Open 3	*
pParent	p271606	Parent child groups offer: Open 4	*
pParent	p271607	Parent child groups offer: Open 5	*
pParent	p401001	Other reasons behind relocation to Germany	*
pParent	p403041	Other reasons behind partner's relocation to Germany	*
pParent	p435030	Religion and religiosity: other religious community	*
pParent	p524410	Nature of health impairment: Other	*
pParent	p524428	Type of health impairment: other	*
pParent	pa10002	ELFRA2: syntax item 2	*
pParent	pa10048	ELFRA2: syntax item 48	*
Weights	px80101	Sample: Federal State	
pParent	p751001_g2	Place of Residence (RS Federal State)	
pParent	p529000	Weight at birth	
pParent	p529001	Size at birth	
pParent	pa01150	Facility fees: Euro	
pParent	pa01250	Childminder fees: Euros	
pParent	p731911	Number of employees respondent	
pParent	p731961	Number of employees partner	
pEducator	e209105	Day care institution: Duration of professional activity: Current institution	*
pEducator	e209111	Day care institution Core group: Room use: Size of rooms	*
pParent	p435030_g1	Religion and religiosity: other religious community	
pParent	pa01151	Facility fees: Cents	
pParent	pa01251	Childminder fees: Cents	
pParentMicrom			*
pEducator			
pEducatorChildminder			

A.2 Recoding of countries/languages/nationalities

The following tables summarize the complete recoding procedure of countries and languages/nationalities. Each category of the aggregated version (**bold**) is followed by all entries which have been subsumed to it. All entries are sorted according to their numeric code.

Table 8: Complete recoding map of all countries

country	
-23	parent not known
-21	Germany and foreign country, not codeable
-20	Foreign country, but not codeable [9005] Kurdish areas, [9991] German-speaking country abroad, [9992] English-speaking country abroad, [9993] Spanish-speaking country abroad
1	Germany [0] Germany
2	Italy [137] Italy
3	Poland [152] Poland
4	Romania [154] Romania
5	Turkey [163] Turkey
6	Former Yugoslavia [120] Yugoslavia, [122] Bosnia and Herzegovina, [130] Croatia, [131] Slovenia, [132] Serbia and Montenegro, [138] Yugoslavia, Federal Republic of, [140] Montenegro, [144] Macedonia, [150] Kosovo, [170] Serbia
7	Former Soviet Union [127] Estonia, [139] Latvia, [142] Lithuania, [146] Moldova, Republic of, [159] Soviet Union, [160] Russian Federation, [166] Ukraine, [169] Belarus, [422] Armenia, [425] Azerbaijan, [430] Georgia, [444] Kazakhstan, [450] Kyrgyzstan, [470] Tajikistan, [471] Turkmenistan, [477] Uzbekistan
8	Central and South America, Caribbean [320] Antigua and Barbuda, [322] Barbados, [323] Argentina, [324] Bahamas, [326] Bolivia, Plurinational State of, [327] Brazil, [328] Guyana, [330] Belize, [332] Chile, [333] Dominica, [334] Costa Rica, [335] Dominican Republic, [336] Ecuador, [337] El Salvador, [340] Grenada, [345] Guatemala, [346] Haiti, [347] Honduras, [349] Colombia, [351] Cuba, [353] Mexico, [354] Nicaragua, [355] Jamaica, [357] Panama, [359] Paraguay, [361] Peru, [364] Suriname, [365] Uruguay, [366] Saint Lucia, [367] Venezuela, Bolivarian Republic of, [369] Saint Vincent and the Grenadines, [370] Saint Kitts and Nevis, [371] Trinidad and Tobago
9	Northern and Western Europe [124] Belgium, [126] Denmark, [128] Finland, [129] France, [135] Ireland, [136] Iceland, [141] Liechtenstein, [143] Luxembourg, [147] Monaco, [148] Netherlands, [149] Norway, [151] Austria, [157] Sweden, [158] Switzerland, [168] United Kingdom, [193] Norwegian dependent territories in Europe, [195] British dependent territories in Europe, [196] Danish dependent territories in Europe, [295] British dependent territories in Africa, [298] French dependent territories in Africa, [392] Dutch dependent territories in America, [395] British dependent territories in America, [396] Danish dependent territories in America, [398] French dependent territories in America, [495] British dependent territories in Asia, [593] Norwegian dependent territories in Australia, Oceania, Antarctica, [595] British dependent territories in Australia, Oceania, Antarctica, [598] French dependent territories in Australia, Oceania, Antarctica

(...)

Table 8: (continued)

	country
10	North America [348] Canada, [368] United States, [391] US dependent territories in America, [591] US dependent territories in Australia, Oceania, Antarctica
11	Oceania/Polynesia [523] Australia, [524] Solomon Islands, [526] Fiji, [527] Cook Islands, [530] Kiribati, [531] Nauru, [532] Vanuatu, [533] Niue, [536] New Zealand, [537] Palau, [538] Papua New Guinea, [540] Tuvalu, [541] Tonga, [543] Samoa, [544] Marshall Islands, [545] Micronesia, Federated States of, [590] New Zealand's dependent territories in Australia, Oceania, Antarctica, [594] Australian dependent territories in Australia, Oceania, Antarctica
12	Other Middle East and North Africa [221] Algeria, [248] Libya, [252] Morocco, [285] Tunisia, [287] Egypt, [421] Yemen, [424] Bahrain, [438] Iraq, [439] Iran, Islamic Republic of, [441] Israel, [445] Jordan, [447] Qatar, [448] Kuwait, [451] Lebanon, [456] Oman, [469] United Arab Emirates, [472] Saudi Arabia, [475] Syrian Arab Republic, [499] other Asia
13	Other Africa [223] Angola, [224] Eritrea, [225] Ethiopia, [226] Lesotho, [227] Botswana, [229] Benin, [230] Djibouti, [231] Côte d'Ivoire, [232] Nigeria, [233] Zimbabwe, [236] Gabon, [237] Gambia, [238] Ghana, [239] Mauritania, [242] Cape Verde, [243] Kenya, [244] Comoros, [245] Congo, [246] Congo, the Democratic Republic of the, [247] Liberia, [249] Madagascar, [251] Mali, [253] Mauritius, [254] Mozambique, [255] Niger, [256] Malawi, [257] Zambia, [258] Burkina Faso, [259] Guinea-Bissau, [261] Guinea, [262] Cameroon, [263] South Africa, [265] Rwanda, [267] Namibia, [268] Sao Tome and Principe, [269] Senegal, [271] Seychelles, [272] Sierra Leone, [273] Somalia, [274] Equatorial Guinea, [276] Sudan, [281] Swaziland, [282] Tanzania, United Republic of, [283] Togo, [284] Chad, [286] Uganda, [289] Central African Republic, [291] Burundi
14	Other Asia [423] Afghanistan, [426] Bhutan, [427] Myanmar, [429] Brunei Darussalam, [431] Sri Lanka, [432] Vietnam, [434] Korea, Democratic People's Republic of, [436] India, [437] Indonesia, [442] Japan, [446] Cambodia, [449] Lao People's Democratic Republic of, [454] Maldives, [457] Mongolia, [458] Nepal, [460] Bangladesh, [461] Pakistan, [462] Philippines, [465] Taiwan, Province of China, [467] Korea, Republic of, [474] Singapore, [476] Thailand, [479] China, [482] Malaysia, [483] Timor-Leste, Democratic Republic of
15	Other Central and Eastern Europe [125] Bulgaria, [155] Slovakia, [162] Czechoslovakia, [164] Czech Republic, [165] Hungary
16	Other Southern Europe [121] Albania, [123] Andorra, [134] Greece, [145] Malta, [153] Portugal, [156] San Marino, [161] Spain, [167] Holy See (Vatican City State), [181] Cyprus, [297] Spanish dependent territories in Africa
17	Other [996] unknown foreign countries

Table 9: Complete recoding map of all languages/nationalities

	language/nationality
1	Arabic [1011801] Arabic
2	Bosnian [1021519] Bosnian
3	Bulgarian [1022112] Bulgarian
4	German [1070518] German, [1140419] Low German; Low Saxon; German, Low; Saxon, Low
5	English [1051407] English
6	French [1061805] French
7	Greek [1071803] Greek, Ancient (to 1453), [1071805] Greek, Modern (1453-)
8	Italian [1092001] Italian
9	Croatian [1081822] Croatian
10	Kurdish [1112118] Kurdish
11	Polish [1161512] Polish
12	Portuguese [1161518] Portuguese
13	Russian [1182119] Russian
14	Romanian [1182113] Romanian; Moldavian; Moldovan
15	Serbian [1191816] Serbian
16	Spanish [1191601] Spanish; Castilian
17	Turkish [1202118] Turkish
18	Vietnamese [1220905] Vietnamese
19	Asian languages [1010305] Achinese, [1011111] Akkadian, [1011803] Official Aramaic (700-300 BCE); Imperial Aramaic (700-300 BCE), [1011913] Assamese, [1012205] Avestan, [1012301] Awadhi, [1020112] Baluchi, [1020114] Balinese, [1020514] Bengali, [1020815] Bhojpuri, [1020908] Bihari languages, [1020911] Bikol, [1021801] Braj, [1022011] Batak languages, [1022101] Buriat, [1022107] Buginese, [1022118] Burmese

(...)

Table 9: (continued)

language/nationality

, [1030502] Cebuano, [1030807] Chagatai, [1030809] Chinese, [1030811] Chuukese, [1031303] Chamic languages, [1040125] Land Dayak languages, [1040922] Divehi; Dhivehi; Maldivian, [1041509] Dogri, [1042615] Dzongkha, [1051224] Elamite, [1060912] Filipino; Pilipino, [1070125] Gayo, [1071514] Gondi, [1072110] Gujarati, [1080502] Hebrew, [1080912] Hiligaynon, [1080913] Himachali languages; Western Pahari languages, [1080914] Hindi, [1080920] Hittite, [1081314] Hmong; Mong, [1090201] Iban, [1091215] Iloko, [1091404] Indonesian, [1100122] Javanese, [1101614] Japanese, [1101618] Judeo-Persian, [1101802] Judeo-Arabic, [1110101] Kara-Kalpak, [1110103] Kachin; Jingpho, [1110114] Kannada, [1110118] Karen languages, [1110119] Kashmiri, [1110123] Kawi, [1110126] Kazakh, [1110801] Khasi, [1110813] Central Khmer, [1110815] Khotanese; Sakan, [1110918] Kirghiz; Kyrgyz, [1111511] Konkani, [1111518] Korean, [1111821] Kurukh, [1112113] Kumyk, [1120108] Lahnda, [1120115] Lao, [1120526] Lezghian, [1122119] Lushai, [1130104] Madurese, [1130107] Magahi, [1130109] Maithili, [1130111] Makasar, [1130112] Malayalam, [1130118] Marathi, [1130125] Malay, [1130418] Mandarin, [1130914] Minangkabau, [1131403] Manchu, [1131409] Manipuri, [1131415] Manobo languages, [1131514] Mongolian, [1132114] Munda languages, [1132318] Marwari, [1140516] Nepali, [1140523] Nepal Bhasa; Newari, [1140901] Nias, [1151809] Oriya, [1151919] Ossetian; Ossetic, [1152001] Turkish, Ottoman (1500-1928), [1160107] Pangasinan, [1160113] Pampanga; Kapampangan, [1160114] Panjabi; Punjabi, [1160515] Persian, Old (ca.600-400 B.C.), [1160518] Persian, [1160814] Phoenician, [1161209] Pali, [1161514] Pohnpeian, [1161801] Prakrit languages, [1162119] Pushto; Pashto, [1180110] Rajasthani, [1190108] Yakut, [1190113] Samaritan Aramaic, [1190114] Sanskrit, [1190119] Sasak, [1190120] Santali, [1190512] Selkup, [1190814] Shan, [1190914] Sinhala; Sinhalese, [1191404] Sindhi, [1191507] Sogdian, [1192114] Sundanese, [1192124] Sumerian, [1192503] Classical Syriac, [1200113] Tamil, [1200120] Tatar, [1200512] Telugu, [1200520] Tetum, [1200711] Tajik, [1200712] Tagalog, [1200801] Thai, [1200902] Tibetan, [1202111] Turkmen, [1202522] Tuvian, [1210701] Ugaritic, [1210907] Uighur; Uyghur, [1211804] Urdu, [1212602] Uzbek, [1230118] Waray, [1250116] Yapese, [1260801] Zhuang; Chuang

20

African languages

[1010118] Afar, [1010308] Acoli, [1010401] Adangme, [1010608] Afrihili, [1010618] Afrikaans, [1011101] Akan, [1011308] Amharic, [1020104] Banda languages, [1020109] Bamileke languages, [1020113] Bambara, [1020119] Basa, [1020510] Beja; Bedawiyet, [1020513] Bemba, [1020518] Berber languages, [1020914] Bini; Edo, [1021420] Bantu (Other), [1031516] Coptic, [1040914] Dinka, [1042101] Duala, [1042521] Dyula, [1050609] Efik, [1050725] Egyptian (Ancient), [1051101] Ekajuk, [1052305] Ewe, [1052315] Ewondo, [1060114] Fang, [1060120] Fanti, [1061514] Fon, [1062112] Fulah, [1070101] Ga, [1070201] Gbaya, [1070526] Geez, [1071802] Grebo, [1080121] Hausa, [1080518] Herero, [1090215] Igbo, [1091015] Ijo languages, [1091405] Indo-European languages, [1110102] Kabyle, [1110113] Kamba, [1110121] Kanuri, [1110911] Kikuyu; Gikuyu, [1110914] Kinyarwanda, [1111302] Kimbundu, [1111514] Kongo, [1111605] Kpelle, [1111815] Kru languages, [1112101] Kuanyama; Kwanyama, [1120113] Lamba, [1120914] Lingala, [1121512] Mongo, [1121526] Lozi, [1122101] Luba-Lulua, [1122102] Luba-Katanga, [1122107] Ganda, [1122114] Lunda, [1122115] Luo (Kenya and Tanzania), [1130114] Mandingo, [1130119] Masai, [1130514] Mende, [1131207] Malagasy, [1131519] Mossi, [1140212] Ndebele, South; South Ndebele, [1140405] Ndebele, North; North Ndebele, [1140415] Ndonga, [1141915] Pedi; Sepedi; Northern Sotho, [1142501] Chichewa; Chewa; Nyanja, [1142513] Nyamwezi, [1142514] Nyankole, [1142515] Nyoro, [1142609] Nzima, [1151813] Oromo, [1182114] Rundi, [1190104] Sandawe, [1190107] Sango, [1190904] Sidamo, [1191201] Slavic languages, [1191401] Shona, [1191411] Soninke, [1191513] Somali, [1191514] Songhai languages, [1191520] Sotho, Southern, [1191818] Serer, [1191923] Swati, [1192111] Sukuma, [1192119] Susu, [1192301] Swahili, [1200513] Timne, [1200907] Tigre, [1200918] Tigrinya, [1200922] Tiv, [1201308] Tamashek, [1201507] Tonga (Nyasa), [1201914] Tswana, [1201915] Tsonga, [1202113] Tumbuka, [1202309] Twi, [1211302] Umbundu, [1220109] Vai, [1220514] Venda, [1230112] Walamo, [1231512] Wolof, [1240815] Xhosa, [1250115] Yao, [1251518] Yoruba, [1260514] Zenaga, [1261404] Zande languages, [1262112] Zulu

(...)

Table 9: (continued)

	language/nationality
21	<p>Other languages</p> <p>[1010211] Abkhazian, [1011202] Albanian, [1011205] Aleut, [1011207] Algonquian languages, [1011407] English, Old (ca.450-1100), [1011601] Apache languages, [1011807] Aragonese, [1011813] Armenian, [1011814] Mapudungun; Mapuche, [1011816] Arapaho, [1011823] Arawak, [1012008] Athapascan languages, [1012119] Australian languages, [1012201] Avaric, [1012513] Aymara, [1012605] Azerbaijani, [1020111] Bashkir, [1020117] Basque, [1020512] Belarusian, [1020919] Bislama, [1021201] Siksika, [1021805] Breton, [1030104] Caddo, [1030118] Galibi Carib, [1030120] Catalan; Valencian, [1030121] Caucasian languages, [1030801] Chamorro, [1030802] Chibcha, [1030805] Chechen, [1030813] Mari, [1030814] Chinook jargon, [1030815] Choctaw, [1030816] Chipewyan; Dene Suline, [1030818] Cherokee, [1030821] Church Slavonic; Old Slavonic; Church Slavonic; Old Bulgarian; Old Church Slavonic, [1030822] Chuvash, [1030825] Cheyenne, [1031518] Cornish, [1031519] Corsican, [1031605] Creoles and pidgins, English based, [1031805] Cree, [1031816] Creoles and pidgins, [1032605] Czech, [1040111] Dakota, [1040114] Danish, [1040512] Delaware, [1040718] Dogrib, [1042113] Dutch, Middle (ca.1050-1350), [1042120] Dutch; Flemish, [1051413] English, Middle (1100-1500), [1051615] Esperanto, [1051920] Estonian, [1060115] Faroese, [1060910] Fijian, [1060914] Finnish, [1061813] French, Middle (ca.1400-1600), [1061815] French, Old (842-ca.1400), [1061825] Western Frisian, [1062118] Friulian, [1070515] Georgian, [1070912] Gilbertese, [1071201] Gaelic; Scottish Gaelic, [1071205] Irish, [1071207] Galician, [1071222] Manx, [1071308] German, Middle High (ca.1050-1500), [1071508] German, Old High (ca.750-1050), [1071518] Gorontalo, [1071520] Gothic, [1071814] Guarani, [1072309] Gwich'in, [1080109] Haida, [1080120] Haitian; Haitian Creole, [1080123] Hawaiian, [1081315] Hiri Motu, [1082114] Hungarian, [1082116] Hupa, [1090305] Icelandic, [1090415] Ido, [1091121] Inuktitut, [1091205] Interlingue; Occidental, [1091401] Interlingua (International Auxiliary Language Association), [1091611] Inupiaq, [1091815] Iroquoian languages, [1110112] Kalaallisut; Greenlandic, [1111513] Komi, [1111519] Kosraean, [1112120] Kutenai, [1120104] Ladino, [1120120] Latin, [1120122] Latvian, [1120913] Limburgan; Limburger; Limburgish, [1120920] Lithuanian, [1122026] Luxembourgish; Letzeburgesch, [1122109] Luiseno, [1130103] Macedonian, [1130108] Marshallese, [1130115] Maori, [1130701] Irish, Middle (900-1200), [1130903] Mi'kmaq; Micmac, [1131220] Maltese, [1131508] Mohawk, [1132119] Creek, [1132514] Mayan languages, [1140108] Nahuatl languages, [1140121] Nauru, [1140122] Navajo; Navaho, [1140903] Niger-Kordofanian languages, [1140921] Niuean, [1141415] Norwegian Nynorsk; Nynorsk, Norwegian, [1141502] Bokmål, Norwegian; Norwegian Bokmål, [1141514] Norse, Old, [1141518] Norwegian, [1150309] Occitan (post 1500); Provençal, [1151009] Ojibwa, [1151901] Osage, [1152015] Otomian languages, [1160112] Pahlavi, [1160116] Papiamentu, [1160121] Palauan, [1161815] Provençal, Old (to 1500), [1172105] Quechua, [1180116] Rapanui, [1180118] Rarotongan; Cook Islands Maori, [1181508] Romansch, [1181513] Romany, [1190112] Salishan languages, [1190315] Scots, [1190701] Irish, Old (to 900), [1190915] Siouan languages, [1191215] Slovak, [1191222] Slovenian, [1191305] Northern Sami, [1191309] Sami languages, [1191315] Samoan, [1191804] Sardinian, [1192305] Swedish, [1200108] Tahitian, [1200518] Tereno, [1201112] Tokelau, [1201209] Tlingit, [1201514] Tonga (Tonga Islands), [1201609] Tok Pisin, [1201909] Tsimshian, [1202120] Altaic languages, [1202212] Tuvalu, [1211118] Ukrainian, [1221512] Volapük, [1221520] Votic, [1230111] Wakashan languages, [1230119] Washo, [1230512] Welsh, [1230514] Sorbian languages, [1231214] Walloon, [1250904] Yiddish, [1251611] Yupik languages, [1260116] Zapotec, [1262114] Zuni, [1262601] Zaza; Dimili; Dimli; Kirdki; Kirmanjki; Zazaki</p>

B References

- Antoni, M., Drasch, K., Kleinert, C., Matthes, B., Ruland, M., & Trahms, A. (2010). *Arbeiten und Lernen im Wandel. Teil I: Überblick über die Studie.*
- Hundepool, A., Domingo-Ferrer, J., Franconi, L., Giessing, S., Nordholt, E. S., Spicer, K., & de Wolf, P.-P. (2012). *Statistical Disclosure Control.* Chichester: Wiley.
- Koberg, T. (2016). Disclosing the National Educational Panel Study. In H.-P. Blossfeld, J. von Maurice, M. Bayer, & J. Skopek (Eds.), *Methodological Issues of Longitudinal Surveys. The Example of the National Educational Panel Study (in press).* Wiesbaden: Springer VS.
- Lane, J., Heus, P., & Mulcahy, T. (2008). Data Access in a Cyber World: Making Use of Cyber-infrastructure. *Transactions on Data Privacy, 1*, 2–16.
- Müller, W., Blien, U., & Knoche, P. (1991). Die faktische Anonymität von Mikrodaten. *Schriftenreihe Forum der Bundesstatistik, 19.*