



Access to and specifics of detailed national LFS data – the case of Slovenia

Sebastian Kočar
Social Science Data Archives
University of Ljubljana

4th DwB workshop,
Ljubljana, September 17th 2014

The content of the presentation

- How to access LFS microdata in Slovenia?
- What microdata are available to different types of users?
- EU-LFS and Slovenian LFS (ADS survey) differences
- Preparing LFS microdata
- Preparing metadata for LFS microdata
- Distribution of LFS microdata and metadata
- Promotion of LFS microdata use
- Plans for the future

Access to LFS microdata in Slovenia

- Microdata available to registered researchers/PhD students working with registered researchers
- SORS research data lab
- SORS remote access connection (large secure public administration network)
- Anonymised data (EU LFS equivalent protection) on CDs
- The process is fairly quick and simple, takes approximately 1-2 months (Data Protection Committee)
- PUF versions are available on ADP website for free (simple online registration, takes 1 day)

LFS microdata for different types of users

<i>Type of users/microdata</i>	Deindividualized LFS microdata	Anonymised LFS microdata (SUF)	Public Use LFS microdata (PUF)
Registered researchers	Research data lab/remote access	CDs	Available on ADP website (simple registration needed)
PhD students	Research data lab/remote access	CDs	
Students	No access	No access	
Public	No access	No access	

EU-LFS and Slovenian LFS (ADS survey) differences

- Variables (additional variables in SORS databases)
- The level of protection (EU anonymisation criteria)
 - Aggregation (e.g. 5-year bands)
 - Top- and low- coding
 - Numeric data -> Categorical data - Deciles (e.g. income)
- Time series distributed (SORS series starts in 1995)

Preparing deindividualized LFS microdata

- preparing deindividualized microdata in the safe room environment
- SPSS is used, SPSS syntax is written
- variable and value labels, missing values are added to the dataset; additional logical control is made, unneeded variables are deleted, variables in databases are connected to codebooks used
- by using SPSS syntax prepared, microdata can be exported in any desired format, readable by variety of software used by researchers

Preparing deindividualized microdata - sources

adsi1_g1_v1_v2 - Notepad

id	stifra stifra	okLIC	tip	datum	starost	ank	refcted	v1_1	v1_4	v1_5
1	1	1	1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	2	2	2

adsi1_om1_v1_v2.gps - IBM SPSS Statistics Syntax Editor

```

GET DATA
  /TYPE=TEXT
  /FILES='adsi11_g1_v1_v1.txt'
EXECUTE.

GET DATA
  /TYPE=TEXT
  /FILES='adsi11_g1_v1_v1.txt'
EXECUTE.

DELET VAR IN(1) F1_2
DELET VAR IN(1) F1_3
RENAME VAR IN(1) F1_2=F1_2
RENAME VAR IN(1) F1_3=F1_3
COMPUTE
  /METHOD=INCOMPATIBLE
EXECUTE.

DELET VAR IN(1) F1_3
RENAME VAR IN(1) F1_3=F1_3
COMPUTE
  /METHOD=INCOMPATIBLE
EXECUTE.

DELET VAR IN(1) F1_3
RENAME VAR IN(1) F1_3=F1_3
COMPUTE
  /METHOD=INCOMPATIBLE
EXECUTE.

DELET VAR IN(1) F1_3
RENAME VAR IN(1) F1_3=F1_3
COMPUTE
  /METHOD=INCOMPATIBLE
EXECUTE.

DELET VAR IN(1) F1_3
RENAME VAR IN(1) F1_3=F1_3
COMPUTE
  /METHOD=INCOMPATIBLE
EXECUTE.

DELET VAR IN(1) F1_3
RENAME VAR IN(1) F1_3=F1_3
COMPUTE
  /METHOD=INCOMPATIBLE
EXECUTE.

DELET VAR IN(1) F1_3
RENAME VAR IN(1) F1_3=F1_3
COMPUTE
  /METHOD=INCOMPATIBLE
EXECUTE.

DELET VAR IN(1) F1_3
RENAME VAR IN(1) F1_3=F1_3
COMPUTE
  /METHOD=INCOMPATIBLE
EXECUTE.

DELET VAR IN(1) F1_3
RENAME VAR IN(1) F1_3=F1_3
COMPUTE
  /METHOD=INCOMPATIBLE
EXECUTE.
    
```

adsi1_g1_v1_v2.sav [DataSet1] - IBM SPSS Statistics Data Editor

Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align
1 sid	Numeric	8	2	Statistični ident.	None	None	8	Right
2 l1stfra	Numeric	8	2	Val anketiranja	None	None	8	Right
3 sifra	Numeric	8	2	Šifra gospod.	None	None	8	Right
4 k0	Numeric	8	2	Št. član gospo.	None	None	8	Right
5 val	Numeric	8	2	Val anketiranja	(1,00, 1 val)	None	8	Right
6 tip	Numeric	8	2	Tip anketiranja	(2,00, teren...	None	8	Right
7 datum	Numeric	8	2	Datum anketa	None	None	8	Right
8 starost	Numeric	8	2	Starost anketi	None	None	8	Right
9 ank	Numeric	8	2	Šifra anketa	None	None	8	Right
10 rfehd	Numeric	8	2	Referenčni tede.	None	None	8	Right
11 v1_1	Numeric	8	2	Ali bo oseba od.	(1,00, Da)	None	8	Right
12 v1_4	Numeric	8	2	Razmerja med	(1,00, njego...	None	8	Right
13 v1_5	Numeric	8	2	Njegovih par	(1,00, oseb...	None	8	Right
14 v1_6	Numeric	8	2	Njegovih oče	(1,00, otrok ...	None	8	Right
15 v1_7	Numeric	8	2	Njegova/ena	(1,00, otrok ...	None	8	Right
16 v1_8	Numeric	8	2	Spol	(1,00, mešk...	None	8	Right
17 v1_9a	Numeric	8	2	Dan rojstva	None	None	8	Right
18 v1_9b	Numeric	8	2	Mesec rojstva	None	None	8	Right
19 v1_9c	Numeric	8	2	Leto rojstva	None	None	8	Right
20 v1_10	Numeric	8	2	Ali imate sloven.	(1,00, Da)	None	8	Right
21 v1_11a	Numeric	8	2	Državljanstvo k.	None	None	8	Right
22 v1_12	Numeric	8	2	Ali ste bili rojen.	(1,00, Da)	None	8	Right
23 v1_13a	Numeric	8	2	V kateri državi	(2,00, Sle n...	None	8	Right
24 v1_14	Numeric	8	2	Kateroga leta s.	None	None	8	Right
25 v1_15a	Numeric	8	2	Iz katere države	None	None	8	Right
26 v1_16	Numeric	8	2	Ali imate na te	(1,00, stah...	None	8	Right
27 v1_17	Numeric	8	2	Kakšen je vaš	(1,00, Sams...	None	8	Right
28 v2_1	Numeric	8	2	Vaša dosežna.	(1,00, ste br...	None	8	Right
29 v2_2a	Numeric	8	2	Kaj ste po izbor	None	None	8	Right
30 v2_3	Numeric	8	2	Kateroga leta s.	None	None	8	Right
31 v2_4	Numeric	8	2	Ali ste v zadnji	(1,00, Da)	None	8	Right
32 v2_5	Numeric	8	2	Katere stopnje	(1,00, Omo...	None	8	Right
33 v2_6	Numeric	8	2	Ali ste v zadnji	(1,00, Da)	None	8	Right
34 v2_7	Numeric	8	2	Koliko ur ste v	None	None	8	Right
35 v2_8	Numeric	8	2	Tečaj, seminar	(1,00, potre...	None	8	Right



Vprašanje

1.1. Vrtiljak članov gospodinjstva	
1.2. Imena članov gospodinjstva	
SAMO ZA PONOVLJENO ANKETIRANJE	
1.3. Kje živi oseba?	
1.3.1. V kateri državi	
1.3.2. V kateri občini	
1.3.3. V kateri ulici	
1.3.4. V kateri hišni številki	
1.4. Razloge za neodvisno gospodinjstvo	
1.4.1. Sami živijo	
1.4.2. Sami živijo zaradi	
1.4.3. Sami živijo zaradi	
1.4.4. Sami živijo zaradi	
1.4.5. Spol	
1.4.6. Starost	
1.4.7. Vzgoja	
1.4.8. Ali imate slovensko državljanstvo?	
1.4.9. Ali ste v zadnji	
1.4.10. Katerega leta s	
1.4.11. Iz katere države	
1.4.12. Dovoljenje za delo	
1.4.13. Kakšen je vaš zakoniti stan?	
1.4.14. Ali imate na tem naslovu stalno prebivališče?	
1.4.15. Datum rojstva	
1.4.16. V kateri državi ste bili rojeni?	
1.4.17. Katerega leta ste se prvič selili v Slovenijo?	
1.4.18. Ali imate na tem naslovu stalno prebivališče?	
1.4.19. Datum rojstva	
1.4.20. V kateri državi ste bili rojeni?	
1.4.21. Katerega leta ste se prvič selili v Slovenijo?	

Preskok



4th DwB Workshop, Ljubljana, 2014



Preparing LFS Public Use Files – the purpose of it

- the majority of researchers would benefit from a simpler access to moderately anonymised microdata
- undergraduate students/potential researchers are not familiar with SORS microdata (can't access them in the detailed form), so they are not aware of the advantages of using them
- the anonymization procedure keeps as much statistical information intact as possible, data are of sufficient quality to be used for advanced level of research

Preparing LFS Public Use Files – how we do it

- in cooperation with SORS Sector for General Methodology and Standards (following very strict rules)
- anonymisation procedure which follows Eurostat LFS anonymisation criteria (in SPSS) + sampling in R! (using packages *sdcMicro*, *bethel*, *samplecube*) + recalculation of weights
 - + better quality of data
 - time consuming
- anonymisation using μ -ARGUS
 - + simple and fast
 - suppression of values

LFS METADATA – structured metadata for researchers

- DDI 2 standard is used
- study descriptions are being prepared, ADP DDI extended scheme is used – including methodological, file description, data description, publication, other material etc. metadata fields
- all the required/useful documentation is made available to researchers in one place (codebooks, questionnaires, publications, syntaxes, methodological explanations etc.)
- metadata is being harvested from SORS and EUROSTAT documentation and websites, also by contacting separate SORS departments, responsible for conducting a survey

Distribution of LFS microdata and metadata

- LFS metadata, including descriptive statistics for LFS variables and metadata documentation, are publically available (ADP website)
- HTML browsing document was prepared (for research data lab, remote access)
- PDF study descriptions (DDI standard) are available to all researchers (research data lab, remote access)
- Easy-to-use microdata are already stored on the research data lab hard disk (protected), could be easily transferred to researcher's folder

Promotion of LFS microdata use

- Microdata are under-used, research potential
- Mailing lists
- DwB project
- International conferences and workshops
- National workshops
 - For students
 - For researchers

Plans for the future

- Continuation of the work done (preparing LFS 2012 and 2013 microdata and metadata)
- Including ad-hoc modules
- LFS PUFs
- Selection of key standard socio-demographic concepts, writing routines
- Workshops for microdata users
- Preparing microdata and metadata for other surveys



Thank you for your attention!

Sebastian Kočar

sebastian.kocar@fdv.uni-lj.si

<http://www.adp.fdv.uni-lj.si/>

4th DwB Workshop

Ljubljana, September 17th 2014