

**UNIVERZA V LJUBLJANI  
FAKULTETA ZA DRUŽBENE VEDE**

**Katja Javoršek**

**UTEŽEVANJE V EVROPSKI DRUŽBOSLOVNI  
RAZISKAVI (ESS)**

Diplomsko delo

**Ljubljana, 2008**

**UNIVERZA V LJUBLJANI  
FAKULTETA ZA DRUŽBENE VEDE**

**Katja Javoršek**

Mentor: red. prof. doc. dr. Vasja Vehovar  
Somentorica: doc. dr. Brina Malnar

**UTEŽEVANJE V EVROPSKI DRUŽBOSLOVNI  
RAZISKAVI (ESS)**

Diplomsko delo

**Ljubljana, 2008**

## ZAHVALA

Vse stvari so težke, preden postanejo lahke.  
*arabska modrost*

Zahvaljujem se mentorju doc. dr. Vasji Vehovar za spodbudo in nasvete pri izdelavi diplomskega dela ter somentorici doc. dr. Brini Malnar za napotke s področja ESS.

Hvala staršem in sorodstvu, ki so me spodbujali in financirali med študijem.

Hvala vsem najbližnjim sošolcem in sotrpinom SOC-DI, ki ste mi na tak ali drugačen način pomagali med študijem. Posebej hvala Ani in Meti, brez vaju diplomska ne bi bila takšna, kot je.

Največja zahvala pa gre Mitju za vso moralno podporo in spodbudo.

## **UTEŽEVANJE V EVROPSKI DRUŽBOSLOVNI RAZISKAVI (ESS)**

Diplomsko delo najprej predstavlja mednarodno raziskavo ESS, njeno metodologijo in vzorčenje. Kot vsaka anketa se tudi ESS ne more izogniti problemu pokritja in neodgovorov. Pristranskost, ki pri tem nastane, se lahko odpravi s populacijskim uteževanjem, in sicer poznamo dve osnovni metodi: poststratifikacijo in raking. Pridobivanje populacijskih podatkov, potrebnih za uteževanje, ni najbolj enostavno, v delu primerjam dva vira, prilogo ESS in Eurostatovo LFS. Predstavljene so štiri kontrolne spremenljivke: spol, starost, izobrazba in regija. Zaradi obsežnosti ESS se osredotočim na primerjavo in analizo znotraj treh držav: Belgije, Slovenije in Švedske. V prvem delu empirične analize ugotovim, da so za uteževanje boljši LFS-populacijski podatki in metoda rakinga. Pred začetkom uteževanja pa je potrebno na podatkih preveriti še neodgovore spremenljivk, s katerimi se lahko soočimo na populaciji ali vzorcu. Manjkajoče vrednosti je potrebno primerno obravnavati, kar analiziram v drugem poglavju empirične analize. Predstavim več situacij, s katerimi se soočimo, in prikažem, kako jih obravnavamo. Manjkajoče podatke na populaciji lahko prerazporedimo ali jih ignoriramo, odstotek tistih na vzorcu pa vstavimo v populacijsko tabelo in ostale vrednosti ustrezno preračunamo. Na koncu glede na priejeno populacijsko tabelo utežim vzorce in izračunam faktor povečanja variance. Uteževanje je ustreznejše tam, kjer imamo manjši odstotek manjkajočih vrednosti.

**Ključne besede:** *Evropska družboslovna raziskava, neodgovori, uteževanje, populacijski podatki, manjkajoče vrednosti.*

## **WEIGHTING IN EUROPEAN SOCIAL SURVEY (ESS)**

This thesis presents the cross-national survey ESS, its rigorous methodology and sampling. As every survey, ESS faces the problems of noncoverage and nonresponse, which lead to nonresponse bias. This can be eliminated by population weighting, using poststratification or raking method. The gathering of necessary population data, needed for the weighting, is not an easy task. I compare two sources, ESS's Appendix and Eurostat's LFS. I also discuss four control variables: gender, age, education and region. Because of the extent of the ESS I focus on three countries: Belgium, Slovenia and Sweden. The first part of my empirical analysis assesses the suitability of the population data, where I find out that LFS data and the raking method are more suitable for weighting. Before the weighting starts item nonresponse has to be checked. I present the possible situations and show how the missing values can be dealt with. One can rearrange or ignore these in population data, inserting those on sample in the population table and suitable recalculate the population data. After having weighted the sample by modified population table I calculate the variance inflation factor. Weighting is more appropriate where we have lower percentage of missing values.

**Key words:** *European social survey, nonresponse, weighting, population data, missing values.*

# Kazalo

<b>1</b>	<b><u>UVOD</u></b>	<b>9</b>
1.1	RAZISKOVALNI CILJ IN HIPOTEZE	10
1.2	UPORABLJENA METODOLOGIJA IN STRUKTURA NALOGE	10
<b>2</b>	<b><u>EVROPSKA DRUŽBOSLOVNA RAZISKAVA (ESS)</u></b>	<b>11</b>
2.1	ESS-VPRAŠALNIK	13
2.2	VZORČENJE V ESS	14
<b>3</b>	<b><u>PROBLEM NEPOKRITJA IN NEODGOVOROV</u></b>	<b>16</b>
3.1	VRSTE IN STOPNJE NEODGOVOROV	17
3.2	NEODGOVORI ZNOTRAJ ESS	18
<b>4</b>	<b><u>UTEŽEVANJE</u></b>	<b>21</b>
4.1	VRSTE UTEŽEVANJA	22
4.2	POSTSTRATIFIKACIJA IN RAKING	24
4.3	OSNOVNO UTEŽEVANJE V ESS	26
4.4	UTEŽEVANJE PRI PRVI IN DRUGI MERITVI	27
<b>5</b>	<b><u>PODATKOVNI VIRI</u></b>	<b>28</b>
5.1	ESS-DOKUMENTACIJA	30
5.2	EUROSTAT: LABOUR FORCE SURVEY (LFS)	31
5.3	SPREMENLJIVKE	32
5.3.1	SPOL	32
5.3.2	STAROST	33
5.3.3	IZOBRAZBA	33
5.3.4	REGIJA	37
5.4	PRIPRAVA PODATKOV	40

<b>6 EMPIRIČNA ANALIZA</b>	<b>44</b>
<b>6.1 POPULACIJSKI PODATKI ESS VS. LFS</b>	<b>44</b>
6.1.1 BELGIJA	44
6.1.2 SLOVENIJA	51
6.1.3 ŠVEDSKA	55
<b>6.2 PROBLEM MANJKAOČIH VREDNOSTI V PODATKIH ZA UTEŽEVANJE</b>	<b>57</b>
6.2.1 BELGIJA R1	60
6.2.2 BELGIJA R3	62
6.2.3 SLOVENIJA R2	65
6.2.4 SLOVENIJA R3	66
6.2.5 ŠVEDSKA R3	68
<b>6.3 POTEK UTEŽEVANJA IN ANALIZA UTEŽENIH PODATKOV</b>	<b>72</b>
<b>7 ZAKLJUČEK</b>	<b>77</b>
<b>8 VIRI IN LITERATURA</b>	<b>80</b>
<b>PRILOGE</b>	<b>85</b>
<b>PRILOGA A: KATEGORIZACIJA NUTS ZA VSE TRI OBRAVNAVANE DRŽAVE</b>	<b>85</b>
<b>PRILOGA B: LFS-PODATKI ZA VSE TRI DRŽAVE ZA 2002, 2004 IN 2006</b>	<b>88</b>
B1. BELGIJA	88
B2. SLOVENIJA	94
B3. ŠVEDSKA	95
<b>PRILOGA C: PRIMER SINTAKSE ZA RAKING, BELGIJA R3</b>	<b>100</b>

## Kazalo tabel:

Tabela 3.2.1: Stopnje odgovorov v sodelujočih državah ESS v %	19
Tabela 5.3.3.1: Izobrazbene strukture za BE, SI in SE	36
Tabela 5.3.4.1: Ravni NUTS in število prebivalcev	38
Tabela 5.3.4.2: Primerjava regionalnih kategorij za vse tri meritve	39
Tabela 5.3.4.1: Slovenija, LFS 2004	40
Tabela 5.3.4.2: Združeni LFS-podatki: Slovenija 2004 – spol/starost/izobrazba	41
Tabela 5.3.4.3: Tabela 5.3.4.2 projicirana na 2000 enot (namišljeni vzorec)	41
Tabela 5.3.4.4: Slovenija, LFS 2004	42
Tabela 5.3.4.5: Združeni LFS-podatki: Slovenija 2004 – regija/starost/spol/izobrazba	43
Tabela 5.3.4.6: Tabela 5.3.4.5 projicirana na 2000 enot (izmišljeni vzorec)	43
Tabela 6.1.1.1: Populacijski podatki: Belgija R1 – spol/starost	44
Tabela 6.1.1.2: Populacijski podatki: Belgija R1 spol/izobrazba	45
Tabela 6.1.1.3: Združeni podatki iz ESS-priloge, primerni za raking (spol/izobrazba)	45
Tabela 6.1.1.4: Združeni podatki iz ESS, Belgija R1, primerni za raking (spol/starost)	45
Tabela 6.1.1.5: Združeni podatki iz ESS, Belgija R1, primerni za raking (regija)	46
Tabela 6.1.1.6: Združeni LFS-podatki: Belgija 2002 – spol/starost/izobrazba	46
Tabela 6.1.1.7: Združeni LFS-podatki: Belgija 2002 – regija	46
Tabela 6.1.1.8: Populacija: Belgija R2 spol/starost	47
Tabela 6.1.1.9: Populacija: Belgija R2 izobrazba	47
Tabela 6.1.1.10: Populacija: Belgija R2 regija samo NUTS1	47
Tabela 6.1.1.11: Združeni podatki iz ESS, Belgija R2, primerni za raking (spol/starost)	48
Tabela 6.1.1.12: Združeni podatki iz ESS, Belgija R2, primerni za raking (izobrazba)	48
Tabela 6.1.1.13: Združeni LFS-podatki: Belgija 2004 – spol/starost/izobrazba	48
Tabela 6.1.1.14: Združeni LFS-podatki: Belgija 2004 – regija	48
Tabela 6.1.1.15: Populacija: Belgija R3 spol/starost	49
Tabela 6.1.1.16: Populacija: Belgija R3 izobrazba	49
Tabela 6.1.1.17: Populacija: Belgija R3 regija, samo NUTS1 popravljena	49
Tabela 6.1.1.18: Združeni podatki iz ESS-priloge, primerni za raking (spol/starost)	49
Tabela 6.1.1.19: Združeni podatki iz ESS-priloge, primerni za raking (izobrazba)	50
Tabela 6.1.1.20: Združeni LFS-podatki: Belgija 2006 – spol/starost/izobrazba	50
Tabela 6.1.1.21: Združeni LFS-podatki: Belgija 2006 – regija	50
Tabela 6.1.1.22: Primerjava ESS in LFS (%), skupine za spol, starost in izobrazbo, Belgija	50
Tabela 6.1.2.1: Populacijski podatki: Slovenija – spol/starost/izobrazba	51
Tabela 6.1.2.2: Populacijski podatki: Slovenija – izobrazba/spol/regija	52
Tabela 6.1.2.3: Združeni ESS-podatki: Slovenija 2002 – spol/starost/izobrazba	53
Tabela 6.1.2.4: Združeni LFS-podatki: Slovenija 2002 – spol/starost/izobrazba	53
Tabela 6.1.2.5: Združeni LFS-podatki: Slovenija 2002 – regija	54
Tabela 6.1.2.6: Združeni LFS-podatki: Slovenija 2004 – spol/starost/izobrazba	54
Tabela 6.1.2.7: Združeni LFS-podatki: Slovenija 2004 – regija	54
Tabela 6.1.2.8: Združeni LFS-podatki: Slovenija 2006 – spol/starost/izobrazba	54
Tabela 6.1.2.9: Združeni LFS-podatki: Slovenija 2006 – regija	54
Tabela 6.1.2.10: Primerjava ESS in LFS (%), skupine za spol, starost in izobrazbo, Slovenija	55
Tabela 6.1.3.1: Združeni LFS-podatki: Švedska 2002 – spol/starost/izobrazba	55
Tabela 6.1.3.2: Združeni LFS-podatki: Švedska 2002 – regija	56
Tabela 6.1.3.3: Združeni LFS-podatki: Švedska 2004 – spol/starost/izobrazba	56
Tabela 6.1.3.4: Združeni LFS-podatki: Švedska 2004 – regija	56

<i>Tabela 6.1.3.5: Združeni LFS-podatki: Švedska 2006 – spol/starost/izobrazba</i>	56
<i>Tabela 6.1.3.6: Združeni LFS-podatki: Švedska 2006 – regija</i>	57
<i>Tabela 6.1.3.7: Primerjava ESS in LFS (%), skupine za spol, starost in izobrazbo, Švedska</i>	57
<i>Tabela 6.1.3.1: Pregled situacij v ESS za Belgijo, Slovenijo in Švedsko</i>	59
<i>Tabela 6.2.1.1: Vzorec spol/starost/izobrazba (n), Belgija R1</i>	60
<i>Tabela 6.2.1.2: Vzorec regija (n), Belgija R1</i>	60
<i>Tabela 6.2.1.3: Populacija spol/starost/izobrazba (%), Belgija 2002</i>	60
<i>Tabela 6.2.1.4: Vzorec spol/starost/izobrazba (%), Belgija R1</i>	61
<i>Tabela 6.2.1.5: Dopolnjena populacija spol/starost/izobrazba (%), Belgija 2002/R1</i>	61
<i>Tabela 6.2.1.6: Prirejena populacija spol/starost/izobrazba (N), Belgija 2002/R1</i>	62
<i>Tabela 6.2.1.7: Prirejena populacija regija (N), Belgija 2002/R1</i>	62
<i>Tabela 6.2.2.1: Vzorec spol/starost/izobrazba (n), Belgija R3</i>	62
<i>Tabela 6.2.2.2: Vzorec regija (n), Belgija R3</i>	63
<i>Tabela 6.2.2.3: Prerazporejanje manjkajočih vrednosti na populaciji, Belgija 2006</i>	63
<i>Tabela 6.2.2.4: Populacija spol/starost/izobrazba brez manjkajočih vrednosti (N), Belgija 2006</i>	63
<i>Tabela 6.2.2.5: Vzorec spol/starost/izobrazba (%), Belgija R3</i>	64
<i>Tabela 6.2.2.6: Populacija spol/starost/izobrazba brez manjkajočih vrednosti (%), Belgija 2006</i>	64
<i>Tabela 6.2.2.7: Dopolnjena populacija spol/starost/izobrazba (%), Belgija 2006/R3</i>	64
<i>Tabela 6.2.2.8: Prirejena populacija spol/starost/izobrazba (N), Belgija 2006/R3</i>	65
<i>Tabela 6.2.3.1: Vzorec spol/starost/izobrazba (n), Slovenija R2</i>	65
<i>Tabela 6.2.3.2: Vzorec regija (n), Slovenija R2</i>	65
<i>Tabela 6.2.3.3: Prirejena populacija spol/starost/izobrazba (N), Slovenija 2004/R2</i>	66
<i>Tabela 6.2.4.1: Vzorec spol/starost/izobrazba (n), Slovenija R3</i>	66
<i>Tabela 6.2.4.2: Vzorec regija (n), Slovenija R3</i>	67
<i>Tabela 6.2.4.3: Vzorec spol/starost/izobrazba (%), Slovenija R3</i>	67
<i>Tabela 6.2.4.4: Populacija spol/starost/izobrazba (%), Slovenija 2006</i>	67
<i>Tabela 6.2.4.5: Dopolnjena populacija spol/starost/izobrazba (%), Slovenija 2006/R3</i>	68
<i>Tabela 6.2.4.6: Prirejena populacija spol/starost/izobrazba (N), Slovenija 2006/R3</i>	68
<i>Tabela 6.2.5.1: Vzorec spol/starost/izobrazba (n), Švedska R3</i>	68
<i>Tabela 6.2.5.2: Vzorec regija (n), Švedska R3</i>	69
<i>Tabela 6.2.5.3: Vzorec spol/starost/izobrazba (%), Švedska R3</i>	69
<i>Tabela 6.2.5.4: Populacija spol/starost/izobrazba (N), Švedska 2006, način 1</i>	70
<i>Tabela 6.2.5.5: Populacija regija (N), Švedska 2006, način 1</i>	70
<i>Tabela 6.2.5.6: Prirejena populacija spol/starost/izobrazba (N), Švedska 2006/R3, način 1</i>	70
<i>Tabela 6.2.5.7: Populacija spol/starost/izobrazba (N), Švedska R3, način 2</i>	71
<i>Tabela 6.2.5.8: Populacija regija (N), Švedska R3, način 2</i>	71
<i>Tabela 6.2.5.9: Prirejena populacija spol/starost/izobrazba (N), Švedska 2006/R3, način 2</i>	72
<i>Tabela 6.2.5.1: Primerjava populacije in vzorca (%), Belgija</i>	73
<i>Tabela 6.2.5.2: Primerjava populacije in vzorca (%), Slovenija</i>	73
<i>Tabela 6.2.5.3: Primerjava populacije in vzorca (%), Švedska</i>	74
<i>Tabela 6.2.5.4: Pregled opisnih statistik končnih uteži za izbrane tri države</i>	75

## Kazalo slik:

<i>Slika 5.3.3.1: Tриje izobrazbeni nivoji v LFS</i>	35
--	----

# 1 Uvod

*Comparative sociology is not a particular branch of sociology; it is sociology itself* (Durkheim v Jowell et al. 2007: 1).

Na vsakem koraku se srečujemo z anketami, pa naj bo to samo kratka anonimna anketa preko spleta, dolga telefonska anketa, pisna anketa, ki nam zagotavlja nagrado, ali pa celo uradna anketa. Ne glede na vrsto ankete sta bolj pomembna njen namen in način analize, ki nas pripeljeta do doseganja cilja. Seveda pot do cilja ni vedno lahka, anketa pa je lahko najbolj učinkovita takrat, ko jo lahko s čim primerjamo. Kot pravijo Jowell et al. (2007: 1) je pomembnost zelo natančnega (*angl. rigorous*) primerjalnega raziskovanja za družbene vede neizpodbitna, saj poleg zanimivih razlik med državami in kulturami pomaga razkriti tudi vidike lastnih držav in kultur, ki bi jih bilo samo z domačimi raziskavami težko ali pa celo nemogoče opaziti.

Ena izmed bolj odmevnih je mednarodna primerjalna raziskava, imenovana Evropska družboslovna raziskava (*angl. European Social Survey*), v nadaljevanju tudi ESS. Čeprav raziskava vključuje natančne verjetnostne vzorce, se med enotami pojavlja problem nepokritja in neodgovorov. Ta se sicer lahko rešuje s pomočjo uteževanja, ki pa lahko po drugi strani ustvari nove probleme.

Pri uteževanju se srečamo s problemom pridobivanja primerljivih populacijskih podatkov in problemom priprave kontrolnih spremenljivk; slednji je lahko zaradi manjkajočih vrednosti zelo kompleksen. V tej diplomski nalogi sem za kontrolne spremenljivke vzela spol, starost, izobrazbo in regijo. Osredotočila se bom na tri države, Belgijo, Slovenijo in Švedsko, ki:

1. bolj nazorno prikažejo razliko med populacijskimi podatki iz dveh virov na eni strani in ESS vzorčnimi podatki na drugi strani, kar bo predstavljeno v šestem poglavju, in
2. zajamejo najbolj pogoste situacije, kjer se soočimo z manjkajočimi podatki.

## **1.1 Raziskovalni cilj in hipoteze**

Namen diplomskega dela je ugotoviti, katere populacijske podatke lahko uporabimo za uteževanje, ter pregledati učinke in probleme pri uteževanju s ključnimi populacijskimi kontrolnimi spremenljivkami (spol, starost, izobrazba in regija), upoštevajoč ob tem seveda tudi osnovno vzorčno utež (*angl. design weight*). Z uteževanjem na kontrolne spremenljivke lahko odstranimo nekaj pristranskosti zaradi neodgovorov.

Osnovni hipotezi sta naslednji: 1) Populacijski podatki iz Eurostatove Ankete o delovni sili (LFS) so bolj primerni za uteževanje kot podatki iz ESS-priloge *Populacijske statistike in ostali podatki*. 2) Pri uteževanju je potrebno vzeti v zakup tudi manjkajoče vrednosti in jih primerno obravnavati.

## **1.2 Uporabljena metodologija in struktura naloge**

Sprva se bom lotila pregleda virov in literature, na podlagi katerih bom predstavila ESS, problem neodgovorov in uteževanje. Nato bom analizirala online podatkovne vire – dokumentacijo in vzorčne baze ESS – ter uradne statistike (Eurostat). Poleg bolj enostavne analize populacijskih in vzorčnih podatkov (primerjava odstotkov) bom pokazala, kako lahko pripravimo tabele za uteževanje, in nato prikazala rezultate uteževanja (statistična analiza).

Struktura naloge bo torej naslednja: v naslednjem poglavju bom predstavila, zakaj je raziskava ESS sploh nastala, kakšni so njeni nameni, cilji in za kaj se uporablja. Sledi poglavje s predstavitvijo problema nepokritja in neodgovorov. Ta se lahko rešuje s pomočjo uteževanja, ki bo poglavitna tema četrtega poglavja. Predstavila bom, kako je uteževanje potekalo v prvih dveh meritvah in do kakšnih problemov je prihajalo. V petem poglavju si bomo pogledali vrste populacijskih podatkov in njihovo pripravo za uteževanje. Šesto poglavje je namenjeno empirični analizi, kjer bom preverjala hipotezi ter predstavila in analizirala rezultate. Za konec sledi še zaključek z glavnimi ugotovitvami in smernicami za nadaljnje raziskovanje.

## **2 Evropska družboslovna raziskava (ESS)**

Zamisel o Evropski družboslovni raziskavi se je sprva porodila leta 1989, ko se je ekipa evropskega projekta »Beliefs in Goverment« lotila problema neprimerljivosti raziskav med različnimi narodi. Namen projekta je bil združiti in interpretirati obstoječe podatke glede sprememb skozi čas o socio-političnih orientacijah evropskih državljanov v različnih državah. Ekipa je imela na voljo veliko virov podatkov (Eurobarometer, Evropske mnenjske ankete, International Social Survey Programe idr.), a je bil obseg za natančno primerjalno analizo med več državami in skozi čas, kljub temu da so te študije formirale ključni material in vir za raziskavo, omejen z diskontinuiteto in notranjimi neskladji, poleg tega so bili stroški precej visoki. To odkritje je bil navdih za Evropsko družboslovno raziskavo (Jowell et al. 2007: 2).

Primarni motiv ESS je zagotoviti učenjakom v Evropi ter tudi drugod redne in zelo natančne nize primerjalnih podatkov kot osnovo za merjenje in analiziranje družbenih sprememb. Doseganje kakovosti podatkov je tako primarna skrb (Jowell et al. 2007: 21). Prav stroga metodologija je temelj za visoko kakovost raziskave. Mohler (2007: 166) opisuje Evropsko družboslovno raziskavo kot eksperiment: »Čeprav uporablja dobro utrjene in visoko kvalitetne anketne metode in postopke, so njeni nekompromisno visoki standardi morda edinstveni za primerjalno vedenjsko anketo v Evropi ali kje drugje.«

ESS je torej mednarodno priznana raziskava, ki pokriva več kot 20 evropskih držav. Zasnovana je po najvišjih standardih (tudi metodoloških) in v celoti financirana s strani Evropske komisije (peti in šesti okvirni program), Evropske znanstvene ustanove in akademskih virov znotraj posameznih držav (European social survey 2007a). Kot je opisano v opisu raziskave (Jowell in Malnar 2004; Jowell in Malnar 2006) je poglavitni cilj Evropske družboslovne raziskave pospeševanje razvoja evropskih primerjalnih družboslovnih znanosti, tako da bi združili metode različnih dimenzij nacionalnih raziskav v Evropi in drugje.

Namen ESS je:

- a) sistematična raziskava stališč, značilnosti in navad evropskih državljanov v povezavi z vrsto pomembnih gospodarskih, družbenih in političnih področij oziroma nadzor in interpretacija sprememb navad in vrednot v Evropi ter
- b) raziskava interakcije med posamezniki in evropskimi institucijami.

Raziskava vključuje natančne verjetnostne vzorce, minimalno 70% stopnjo sodelovanja in natančni zapisnik o protokolih glede prevajanja. Na kratko, trudijo se biti čim bolj natančni in zanesljivi.

Raziskava poteka okvirno vsaki dve leti. Prva meritev (*angl. Round 1*, v nadaljevanju tudi R1), ki je pokrivala 22 držav, je potekala v letih 2002/2003. Druga meritev (*angl. Round 2*, v nadaljevanju tudi R2) je pokrivala 26 držav in potekala v letih 2004/2005, tretja meritev (*angl. Round 3*, v nadaljevanju tudi R3) pa v 2006/2007 in pokrivala 25 držav. Trenutno poteka četrta meritev (*angl. Round 4*), ki predvideva objavo podatkov konec septembra 2009 (European social survey 2007a).

Poleg neobičajne strogosti za primerjalno vedenjsko raziskavo sta dve značilnosti ESS privlačili takojšnjo in široko razvejano zanimanje družboslovnih znanstvenikov, kot navajajo Jowell et al. (2007: 4):

- Prva značilnost je razdelitev ESS-vprašalnika na dve polovici, ena je namenjena ključnim merjenjem, ki so enaka v vsaki meritvi, druga pa vsebuje dva tematska modula. Ta dogovor zagotavlja obstoj kontinuitete med meritvami, po drugi strani pa omogoča fleksibilnost vsebine.
- Druga značilnost je njena politika transparentnosti in odprtih dostopov. Vsi protokoli in metode so takoj dostopni na spletni strani Evropske družboslovne raziskave ([www.europeansocialsurvey.org](http://www.europeansocialsurvey.org)), prav tako so dostopni tudi podatki ([www.ess.nsd.uib.no](http://www.ess.nsd.uib.no)). V letu 2005 je še posebej naraslo zanimanje za ta projekt, ko je ESS dobila prestižno Descartesovo nagrado za komunikacijo v znanosti (*angl. Descartes Prize for Excellence in Collaborative Scientific Research*).

Tu lahko omenim, da se raziskava glede pomembnosti uvršča v »deveti, najvišji rang med primerjalne ali kontinuirane raziskave« (Jowell in Malnar 2004, Jowell in Malnar 2006).

Posebnost Evropske družboslovne raziskave je torej ta, da so vsi podatki in pripadajoča dokumentacija takoj objavljeni v lahko dostopni obliki preko svetovnega spletja, tako da lahko analitiki začnejo z resnim delom na podatkih brez kakršnih koli tehničnih ali administrativnih zamud. Podatki so sočasno dostopni vsem ne glede na njihovo prejšnjo vpleteneost. V splošnem je intelektualna vsebina ESS kolektivni produkt, ki je in še bo implicitno ali eksplizitno pritegnil učenjake iz Evrope in drugod (Kolsrud et al. 2007: 140–142). Tako ima vsak lahko dostop do podatkov predvsem glede dokumentacije in načina izbora vzorca, za podatkovne baze se je potrebno registrirati, registracija pa je brezplačna.

Več o podatkih v kasnejših poglavjih, najprej si poglejmo podrobnejše izgled vprašalnika in potek vzorčenja.

## **2.1 *ESS-vprašalnik***

Eno uro dolga osebna anketa (intervju iz oči v oči) vključuje blok vprašanj o splošnih družbenih in političnih stališčih, ki ostaja nespremenjen skozi leta, ter variabilne module vprašanj posvečene posameznim temam (Jowell in Malnar 2004).

Kot pišeta Jowell in Malnar (2006), je vsebinsko raziskava razdeljena na več sklopov. V prvem delu vprašalnika, ki ostaja enak v vseh meritvah ESS, so naslednji:

- sklop vprašanj o stališčih, ki služijo za opazovanje srednjeročnih in dolgoročnih sprememb;

- sklop kazalcev delovanja ključnih družbenih sistemov (gospodarstvo, izobraževanje, zdravstvo idr.), politične participacije, vrednot, kakovosti življenja ipd.;
- sklop vprašanj o demografskih in socialnih značilnostih. Ta blok vprašanj, ki v analitičnih modelih praviloma vsebuje pojasnjujoče spremenljivke, je »v primerjavi s podobnimi empiričnimi projekti zelo obsežen in s tem potencialno močan vir za oblikovanje modelov pojasnjevanja družbenih trendov in razlik med skupinami« (prav tam). Poleg tradicionalnega nabora demografskih spremenljivk vsebuje tudi kazalce individualno-psiholoških in kulturno-vrednotnih značilnostih posameznika.

Drugi del vprašalnika je rotirajoči modul za posebne vsebine. To je variabilni del vprašalnika, ki vsebuje nabor 30–60 kazalcev iz zaključenih tematskih blokov. Cilj je možnost intenzivne in sistematične analize določenega družbenega področja, pri čemer pokriva predvsem aktualne tematike (prav tam).

Dva dela vprašalnika nam omogočata analizo podatkov skozi čas, hkrati pa tudi osredotočenost na trenutna pereča vprašanja.

## 2.2 Vzorčenje v ESS

*The ESS represents a significant step forward in the control of sample design in multi-nation surveys (Häder in Lynn 2007: 51).*

Da so mednarodne ankete veljavne, morajo biti osnovane na verjetnostnih vzorcih primerljivih nacionalnih populacij, prav tako morajo biti odgovori kontrolirani za primerljivost (Kish v Häder in Lynn 2007: 34). Tudi v nacionalnih anketah velja, da naj bi imel vsak državljan enako možnost izbora v vzorec, da lahko govorimo o reprezentativnem vzorcu (Jowell et al. 2007: 6). Vsaka raziskava se trudi, da bi se čim bolj približala reprezentativnem vzorcu in ESS se je tega lotila na svoj način.

Najprej so se lotili definicije enakovrednosti glede na dve osnovni karakteristiki anketnega vzorca: populacijo, ki jo vzorec predstavlja, in natančnost, s katero lahko zagotovi ocene karakteristik te populacije. Populacijo so definirali kot »vse osebe, starejše od 15 let, ki živijo v zasebnih gospodinjstvih znotraj države, ne glede na narodnost, državljanstvo, jezik ali pravni status« (Häder in Lynn 2007: 34). Vsak vzorec naj bi tako zagotavljal statistično nepristranske cenilke karakteristik populacije. Večini namenov in brez previsokih stroškov naj bi zadoščal enostavni slučajni vzorec (SRS – Simple Random Sample) s 1500 enotami, tako da so se tega poskusili držati v vseh državah. Poudarili so, da je enaka natančnost v vseh državah želen cilj, saj je osnovni cilj mednarodnega raziskovanja omogočiti primerjanje med državami, pa naj bo glede osnovnih opisnih statistik, glede asociacij med spremenljivkami ali glede kompleksnih multivariatnih analiz (prav tam: 34–7).

Cilj vzorčne strategije ESS je načrtovanje in implementacija funkcionalnih in ustreznih vzorčnih strategij v vseh sodelujočih državah. Za to potrebujejo naključni verjetnosti vzorec s primerljivimi cenilkami, ki temeljijo na popolni pokritosti ustrezne populacije starejših od 15 let (European social survey 2007b). Kljub enostavni definiciji populacije pa je bilo v nekaterih državah to presenetljivo težko aplicirati na realnost. Nekatere države niso navajene uporabljati tako nizke starostne meje, druge so imele tudi zgornjo starostno mejo (75 ali 80 let). Tudi geografska pokritost populacije ponekod ni popolna, saj bi bili določeni deli predragi (otoki in province) ali pa tudi prenevarni za intervjuje (Häder in Lynn 2007: 38). Dejansko izvrševanje naključnega verjetnostnega vzorca močno variira med državami, saj je izpolnitev odvisna od dostopnosti vzorcev in drugih spremenljivk. Vsi pa temeljijo na osnovi strogih načel verjetnosti in predstavnosti (European social survey 2007b). V anketah je torej težko doseči popolno pokritost vzorca, saj vedno obstajajo manjkajoče enote in se tako srečujemo s problemom nepokritja ter s problemom neodgovorov.

### 3 Problem nepokritja in neodgovorov

Manjkajoči elementi se lahko pojavijo, kadar je okvir neprimeren (*angl. inadequate*), torej ni namenjen naši ciljni populaciji, ali pa nepopoln (*angl. incomplete*), torej je ustrezan, vendar ne vključuje vseh elementov (Kalton in Vehovar 2001: 68). Prvi označuje primer nepokritja, drugi primer neodgovora.

Neodgovori lahko vplivajo na kvaliteto anketnih statistik. Če imajo nerespondenti drugačne vrednosti na spremenljivkah, ki so komponente statistik, se vrednost statistik na respondentih lahko razlikuje od vrednosti celotnega vzorca. Če je odstopanje sistematične narave, prisotno v vseh ponovitvah raziskave, to imenujemo pristranskost zaradi neodgovorov (Groves et al. 2007: 169). Zato se trudimo, da bi čim bolj zmanjšali to pristranskost.

Pristranskost je sistematična napaka, ki jo povzroča člen v anketni verigi in je prisotna v vseh ponovitvah raziskave. Lahko jo povzroča čustveno postavljeno anketno vprašanje ali anketar, ki deluje na velikem področju in uporablja nestandardne pristope in pri tem odstopa od povprečja. Točnost ocene meri povprečna kvadratna napaka (*angl. mean square error*). Definirana je kot pričakovana vrednost kvadratov razlik med vzorčnimi ocenami in pravo populacijsko vrednostjo (Groves; Groves in Biemer; Blemer in Stokes v Kveder 2001: 6–7).

V praksi nastajajo različna odstopanja od idealnega vzorčnega okvira, Kish (v Kalton in Vehovar 2001: 67) jih je strnil v naslednje štiri skupine:

- manjkajoči elementi (*angl. missing elements*): populacijski elementi, ki niso vključeni v vzorčni okvir, čeprav sodijo v ciljno populacijo;
- skupine (*angl. clusters*): vzorčni okvir vključuje na enem naslovu skupino elementov in ne le posamezen element;
- neustrezni elementi (*angl. uneligible elements*): vzorčni okvir vključuje prazne ali tuje elemente, ki ne pripadajo ciljni populaciji;

- podvojene zapise (angl. *duplicate listings*): populacijski elementi se v vzorčnem okviru pojavljajo večkrat.

Sama se bom osredotočila samo na manjkajoče elemente, znotraj enot, ki so odgovarjale, pa na manjkajoče odgovore spremenljivk.

### **3.1 Vrste in stopnje neodgovorov**

Znotraj konceptualnega okvira ustreznih enot je mogoče neodgovor opredeliti na dveh različnih ravneh, in sicer kot nesodelovanje v anketni raziskavi – neodgovor enote (angl. *unit nonresponse*) ter kot nezmožnost ali nepripravljenost odgovoriti na določeno anketno vprašanje (angl. *item nonresponse*) – imenovan tudi neodgovor spremenljivke (Kveder 2001: 4). V tej nalogi se bom bolj osredotočila na to, kako lahko izboljšamo pristranskost zaradi neodgovora enot in kakšne probleme nam pri tem povzročajo neodgovori spremenljivk.

*Neodgovor spremenljivke se kaže kot manjkajoči podatek v zapisu določenega odgovora (spremenljivke) pri elementu, ki je sicer sodeloval v anketni raziskavi. Stopnjo odgovora spremenljivke (angl. *item response rate*) izražamo z razmerjem med številom elementov, za katere pri obravnavani spremenljivki razpolagamo z odgovorom, ter številom vseh ustreznih elementov, ki so v raziskavi na to vprašanje odgovarjali. Stopnja neodgovora spremenljivke (angl. *item nonresponse rate*) pa je komplementarna mera, ki izraža delež elementov z manjkajočo vrednostjo pri obravnavani spremenljivki med vsemi za to spremenljivko ustreznimi elementi (Kalton in Vehovar 2001: 88).*

Kvaliteta ankete je pogosto merjena tudi glede na njeno stopnjo odgovorov. Zato so te stopnje pri ESS pomembna spremenljivka. Vedno se bodo našli respondenti, ki jih intervjuvanci ne bodo mogli locirati v časovnem okviru raziskave (nekontakti). Spet drugi na anketo ne bodo hoteli odgovarjati (zavrnitve), tretji pa zaradi tehničnih razlogov (npr. bolezen, neznanje jezika) ne bodo mogli sodelovati. Žal lahko še posebej v mednarodnih anketah stopnja neodgovorov vpliva na veljavnost primerjave znotraj držav (Billiet et al. 2007:

113). Zgoraj omenjene tri kategorije nerespondentov so splošne, omenjajo jih tudi Fowler (2002: 40) in Groves et al. (2004: 169–170). Neodgovor je lahko posledica različnih razlogov, čeprav prevladujeta predvsem odsotnost ustreznega anketiranca ter njegova zavestna odločitev za nesodelovanje (Kalton in Vehovar 2001: 80).

### **3.2 *Neodgovori znotraj ESS***

ESS zahteva določene standarde in dokumentacijo, da bi dosegli čim višje stopnje odgovorov, zmanjšali pristranskost in analizirali te pojave na mikro in makro nivoju. Za zvišanje pridobljenih odgovorov so uporabljali več taktik (vsaj štirje obiski anketiranca, izobraževanje anketarjev, nadzor terenskega dela idr.), želena stopnja pa je bila vsaj 70 % (Billiet et al. 2007: 114–115).

Zahteve za vzorec, ki naj bi omogočil posploševanje z vzorca na populacijo, so bile naslednje: verjetnostni vzorec, popolna pokritost ciljne populacije, brez zamenjav, visoke stopnje odgovorov (najmanj 70 %) in najmanj 2000 enot v vzorcu ali vsaj najmanjša učinkovita velikost vzorca (po odpravljenih vzorčnih učinkih) v sodelujočih državah, kar je 1500 ali 800 enot, kjer je manj kot 2 milijona prebivalcev (European social survey 2007c, 2007d: pogl. 2).

Stopnja neodgovorov pa naj ne bi preveč variirala med različnimi populacijskimi skupinami. Takšni problemi vplivajo na nacionalne ankete, prisotnost le-teh pa je še mnogo večja pri mednarodnih anketah. Zaradi širokega spektra problemov se večina mednarodnih anket z njimi ni mogla uspešno spoprijeti. Vsaka mednarodna anketa se brez dvoma sooča s problemi in z različnostmi, saj je nerealno pričakovati, da bo vsaka država uporabila enak način vzorčenja (Jowell et al. 2007: 9).

Tabela 3.2.1 predstavlja stopnje odgovorov v odstotkih, kjer lahko vidimo, da niso vse države dosegle želene stopnje odgovorov, le-ta pa niha med 33.5 % (Švica, R1) in 80 % (Grčija, R1).

**Tabela 3.2.1: Stopnje odgovorov v sodelujočih državah ESS v %**

Meritev:	R1	R2	R3
<b>Država:</b>			
<b>Avstrija</b>	60.4	62.4	64.0
<b>Belgija</b>	59.2	61.2	61.0
<b>Bolgarija</b>	/	/	64.8
<b>Ciper</b>	/	/	67.3
<b>Češka</b>	43.3	55.3	/
<b>Danska</b>	67.6	64.2	50.8
<b>Estonija</b>	/	<b>79.1</b>	65.0
<b>Finska</b>	73.2	70.7	64.4
<b>Francija</b>	43.1	<b>43.6</b>	<b>46.0</b>
<b>Grčija</b>	<b>80.0</b>	78.8	/
<b>Madžarska</b>	69.9	65.9	66.1
<b>Islandija</b>	/	51.3	/
<b>Irska</b>	64.5	62.5	56.8
<b>Izrael</b>	71.0	/	/
<b>Italija</b>	43.7	59.3	/
<b>Latvija</b>	/	/	71.2
<b>Luksemburg</b>	43.9	50.1	/
<b>Nemčija</b>	55.7	51.0	54.5
<b>Nizozemska</b>	67.9	64.3	59.8
<b>Norveška</b>	65.0	66.2	65.5
<b>Poljska</b>	73.2	73.7	70.2
<b>Portugalska</b>	68.8	71.2	<b>72.8</b>
<b>Romunija</b>	/	/	71.8
<b>Rusija</b>	/	/	69.5
<b>Slovaška</b>	/	62.7	73.2
<b>Slovenija</b>	70.5	70.2	65.1
<b>Španija</b>	53.2	54.9	65.9
<b>Švedska</b>	69.5	65.4	65.9
<b>Švica</b>	<b>33.5</b>	48.6	51.5
<b>Turčija</b>	/	50.7	/
<b>Ukrajina</b>	/	66.6	66.4
<b>Velika Britanija</b>	55.5	50.6	54.6

Opomba: Znak / pomeni, da država ni sodelovala v meritvi. Krepko označene številke predstavljajo najvišjo in najnižjo stopnjo v meritvi.

**Vir: ESS DATA 2003b, 2005b, 2007b.**

V tej nalogi se bom osredotočila na tri države, in sicer Belgijo, Slovenijo ter Švedsko, ki so sodelovale v vseh treh meritvah. Belgija je za svoj vzorec uporabila nacionalni register, Slovenija osrednji register populacije in Švedska register populacije. Za vzorčenje je Belgija uporabila enostavno slučajno vzorčenje v mestih, drugje stratifikacijo po skupinah (dve fazи, 324 vzorčnih točk). Slovenija je uporabila stratifikacijo po skupinah (dve fazи, 150 vzorčnih točk), Švedska pa enostavno slučajno vzorčenje (Häder in Lynn 2007: 39, 41).

Vse tri države so približno dosegle in obdržale želeno stopnjo odgovorov, vendar pa se razlikuje stopnja neodgovorov spremenljivk, kar bo podrobnejše predstavljeno kasneje.

V raznih raziskavah neodgovorov so prišli do naslednjih ugotovitev glede vpliva demografskih spremenljivk na neodgovore:

- Groves in Couper, Koch, Groves, Stoop (v Mohler 2007) so naleteli na več sistematičnih razlik med respondenti in nerespondenti ter prišli do soglasja, da v anketah najverjetneje manjkajo mladi zaposleni in upokojeni, ki živijo v enočlanskih gospodinjstvih. Prav tako neradi sodelujejo manj izobraženi ljudje in mestni stanovalci (Mohler 2007: 163).
- Stopnja neodgovorov v osebnih »face-to-face« anketah je pogosto višja znotraj mest kot drugod (Kalton 1983: 66).
- Billiet et al. (2007: 130–131) so analizirali socio-demografske spremenljivke glede pristranskosti zaradi neodgovorov, tudi glede starosti in izobrazbe, vendar niso ugotovili posebne povezanosti.

Demografske spremenljivke naj bi torej vplivale na neodgovore, vendar ni nujno, da povzročajo tudi pristranskost.

Nekateri neodgovori so neizogibni. Poznamo tri pristope k zmanjšanju posledične napake: uporaba »proxy« respondentov, statistična prilagajanja in ponovno anketiranje vzorca nerespondentov. Med statistična prilagajanja štejemo tudi uteževanje (Fowler 2002: 51).

## 4 Uteževanje

Uteževanje je prilagajanje izračunov anketnih statistik, tako da nevtraliziramo škodljive efekte nepokritja, neodgovorov ali neenake verjetnosti izbora v vzorec (Groves et al. 2007: 305). Uteži so uporabne zato, ker v analizi nekaterim vzorčenim elementom pripisujejo večjo relativno pomembnost kot pa drugim, in jih potrebujemo tam, kjer so vzorčne enote izbrane z neenakim verjetnostnim vzorčenjem; prav tako so uporabljeni za poststratifikacijo in pri prilagajanju podatkov za neodgovore (Kalton 1983: 69).

Kalton in Vehovar opozarjata na eno pomembnih in neugodnih posledic uteževanja, tj. povečanje vzorčne variance. Ocena povečanja vzorčne variance, ki nastaja zaradi uporabe uteževanja, izhaja iz elementarnega koeficiente variacije  $CV(w)$  za spremenljivko uteži  $w$ .  $CV^2(w)$  je kvadrat tega koeficiente variacije uteži in izraža razmerje med elementarno varianco uteži in kvadratom aritmetične sredine, kar včasih označujemo tudi z izrazom relativna varianca (angl. relvariance). Pri obravnavi uteževanja ga vključujemo v poseben faktor povečanja variance (angl. *Variance Inflation Factor – VIF*):  $VIF = 1 + CV^2(w)$ , ki izraža povečanje vzorčne variance uteženega vzorca v primerjavi z vzorčno varianco vzorca enake velikosti, kjer uteži niso bile potrebne. Izraz VIF ocenjujemo s cenilko  $vif$ , ki ga lahko izračunamo iz osnovnih opisnih statistik. Če je povprečje uteži enako  $\bar{w} = 1$ , potem ocena za VIF izhaja iz elementarne variance za spremenljivko uteži:

$$vif = 1 + cv(w) = 1 + \frac{sd^2(w)}{\bar{w}_2} = 1 + sd^2(w),$$

kjer je oznaka  $w$  uporabljena v smislu spremenljivke in ne kot vsota uteži (Kalton in Vehovar 2001: 107–110).

VIF se nanaša na fiksne uteži, obstajajo pa tudi variabilne uteži. Le-te so npr. pri populacijskem uteževanju, saj so deleži za vrednosti kontrolnih spremenljivk v primeru enakih in ponavljajočih vzorcev vsakokrat nekoliko drugačni in so tako vsakokrat nekoliko drugačne tudi odgovarjajoče uteži, ki vzorčno strukturo

prilagodijo populacijski. Zato je v vsaki ponovitvi vzorca potrebno ponoviti tudi celoten postopek uteževanja. Na tak način je še teže pravilno ovrednotiti učinke variabilnih uteži, je pa porast v vzorčni varianci zaradi uteževanja manjši kot to izhaja iz VIF in je ocena na osnovi VIF varna (Kalton in Vehovar 2001: 112).

»Koliko je uteževanje v resnici koristno, lahko ocenimo šele z izračunom srednje kvadratne napake (angl. *Mean Squared Error* – *MSE*), ki je vsota kvadrata pristranskosti (angl. *bias*) in variance:  $MSE(\bar{y}) = Bias^2(\bar{y}) + Var(\bar{y})$ « (prav tam: 113). Za izračun srednje kvadratne napake si tako izberemo neko spremenljivko  $y$ , v kar pa se zaradi obsežnosti ne bom spuščala. V empiričnem delu bom predstavila ocene *VIF* za izbrane tri države. Kljub temu pa je potrebno še omeniti, da če se z uteževanjem srednja kvadratna napaka poveča, uporaba uteži verjetno ni smiselna. Tudi Grooves meni, da včasih ni enostavno primerno opraviti statističnih testov, če je anketa sestavljena bolj kompleksno. Uteževanje je za nekatere korelacijske analize lahko celo nepotrebno (Fowler 2002: 27–28). Uteževanje je uspešno le tedaj, ko so nerespondenti glede na ciljno spremenljivko slučajna podmnožica v razredu, kjer želimo opraviti uteževanje. Ker je to v praksi redko, ponavadi uteževanje delno odpravi pristranskost le pri nekaterih spremenljivkah (Kalton in Vehovar 2001: 87).

## 4.1 Vrste uteževanja

Uteževanje se lahko izvaja kot tehnika *stratifikacije po izboru* oziroma poststratifikacija. Pri tej tehniki uporabimo znano porazdelitev populacije za eno ali več dodatnih spremenljivk za izboljšanje natančnosti vzorčnih cenilk. Primer: za populacijo je znana starostna porazdelitev iz zadnjega cenzusa. Vzorec lahko razdelimo na starostne skupine in izračunamo povprečje za spremenljivko  $y$  z vsako starostno skupino ( $y_h$ ). Ta povprečja se lahko skombinirajo v splošno cenliko ( $y_{ps} = W_h y_h$ ), kjer je  $W_h$  delež populacije v starostni skupini  $h$ . Poststratificirano povprečje se lahko označuje tudi kot uteženo povprečje, kjer vsaka enota pridobi utež, sorazmerno z  $N_h/n_h$ . Če ne upoštevamo hipnih problemov neodgovorov in nepokritja, poststratifikacija prilagodi vzorčno

porazdelitev skozi stratume, kar je tudi predmet naključnih fluktuacij, da jo naredi konformno znani populacijski porazdelitvi (Kalton 1983: 74).

Uteževanje vzorca na znano populacijsko porazdelitev ne popravi samo vzorčne fluktuacije, pač pa tudi neodgovore in nepokritja. Npr. če je stopnja neodgovorov višja med mladimi ali če več takšnih manjka v vzorčnem okviru, uteževanje vzorca na znano starostno skupino kompenzira te faktorje. Kljub temu je potrebno vedeti, da je kompenzacija pridobljena glede na dane starostne skupine, in ker lahko obstajajo razlike med posamezniki tudi znotraj teh skupin, bo še vedno prisotne nekaj pristranskosti zaradi neodgovorov (prav tam: 74).

V praksi je lahko pridobivanje uteži zelo zapletena naloga, saj je pogosto potrebna kombinacija prilagajanj. Najprej lahko uteži uporabimo za prilagoditev zaradi neenakih verjetnosti izbora. Nato se te uteži popravi, da se prilagajajo različnim stopnjam neodgovorov znotraj razredov v vzorcu. Končno se popravi še razmerje vzorčne porazdelitve glede na znano populacijsko porazdelitev. Posebna pozornost je potrebna pri razvoju uteži, saj lahko hitro pride do resnih napak (prav tam: 75).

Groves et al. (2007: 322–328) predstavljajo naslednje štiri tipe uteževanja:

- 1) **Uteževanje kot prilagajanje razmerjem v prvi fazi** (*angl. weighting with a First-Stage Ratio Adjustment*). Le-ta se uporablja pri stratificiranem večfaznem vzorcu, s tem uteževanjem pa pridobimo stabilnost cenilk v različnih izborih osnovnih izbranih spremenljivk, tako da je utežena vsota skladna z realizacijo vzorčnega okvirja.
- 2) **Uteževanje zaradi različnih verjetnosti izbora** (*angl. weighting for Differential Selection Probabilities*). Kadar imajo določeni elementi večjo verjetnost izbora v vzorec, npr. da je neka skupina prekomerno predstavljena, se za prilagajanje tega uporabi uteži. V ESS-raziskavi je za to namenjena utež *dweight*.

- 3) **Uteževanje kot prilagajanje na neodgovore enot** (*angl. weighting to adjust for unit nonresponse*). Stopnja neodgovorov ni enaka v vseh populacijskih skupinah, vendar pa je tu predpostavka, da za uteževanje uporabimo tiste spremenljivke, za katere poznamo vrednosti tako za respondentne kot tudi za nerespondente. V praksi morajo biti razredi za prilagajanje omejeni na spremenljivke, katerih vrednosti poznamo pri vseh enotah v vzorcu, npr. starost. Če za enote znotraj vzorčnega razreda predvidevamo, da predstavljajo naključni vzorec vseh vzorčenih oseb, potem odgovori le-teh predstavljajo vzorčno stopnjo, rečemo tudi, da neodgovori manjkajo slučajno (MAR).
- 4) **Poststratifikacijsko uteževanje za zmanjšanje vzorčne variance** kot tudi nepokritja in neodgovorov enot (*angl. poststratification weighting for sampling variance reduction and also undercoverage and unit nonresponse*). Končno uteževanje, ki ga uporabljam v mnogih anketah, je poststratifikacija. Le-ta uporabi zunanje uteži, da vzorčna vsota ustreza zunanji vsoti, ki temelji na ciljni populaciji. Uteževanje je lažje in statistično bolj učinkovito, če je vsota uteži enaka tej zunanji porazdelitvi.

Vse te uteži na koncu se združijo skupaj v končno utež, ki je dodana vzorčni bazi. Nekatere baze lahko vsebujejo poleg končne tudi posamezne komponente. Pri svoji empirični analizi bom uporabila tretji in četrти tip uteževanja, saj so ESS-podatki zbrani tako, da prva dva tipa nista potrebna. Na kratko pa bom predstavila dve metodi uteževanja, in sicer poststratifikacijo in raking.

## 4.2 Poststratifikacija in raking

Za uteževanje vzorčnih podatkov uporabljamo podatke iz zunanjih virov. Da bi lahko primerjali rezultate raziskave za več držav, bi morali podatke utežiti glede na tiste spremenljivke, ki nas v analizi zanimajo. Vendar pa ponavadi nimamo na voljo takšnih podatkov za celotno populacijo, zato utežujemo glede na znane variable, kot so spol, starost, izobrazba in tudi regija. Poznamo dve vrsti populacijskega uteževanja: poststratifikacija in raking.

Za poststratifikacijo potrebujemo popolno tabelo za vse kategorije vseh spremenljivk za uteževanje – po ena celica za vsako skupino. »Primarni cilj poststratifikacije je, da se znebimo potencialne pristranskosti zaradi napak, pridobljenih zaradi odgovorov, nepokritosti in neodgovorov. Sekundarni cilj pa je zmanjšati vzorčne napake« (Holt and Smith in Brick et al. 2003: 1–4).

Poststratifikacija se lahko uporabi tudi takrat, ko smo enote v vzorcu zbrali s disproportionalno stratifikacijo. Sama stratifikacija temelji na predhodni razdelitvi populacije na podpopulacije (stratume), kar se opravi s pomočjo dodatnih populacijskih informacij. V postopku izbiranja vzorca se vzorčenje opravi ločeno in neodvisno za vsak stratum. Velikosti vzorcev v stratumih so tako nadzorovane in pogosto enostavno sorazmerne ustreznim velikostim populacije, kar imenujemo proporcionalna stratifikacija. Pri disproportionalni imajo lahko stratumi različne vzorčne deleže, zato struktura stratumov v vzorcu ne odseva nujno stratumov v populaciji (Kalton in Vehovar 2001: 26).

Za raking metodo potrebujemo manj podatkov, saj so lahko namesto v eni tabeli tudi v več tabelah. Tako potrebujemo manj celic, vsak posameznik pa je vključen v vseh tabelah. Raking je pogosto uporabljen estimacijski postopek, kjer so ceničke nadzorovane glede na robne frekvence populacije. Lahko ga dojemamo kot večdimenzionalni poststratifikacijski postopek, ker so uteži poststratificirane na eno dimenzijo kontrolnih vrednosti, nato so te spremenjene uteži poststratificirane na naslednjo dimenzijo kontrolnih vrednosti. Postopek se nadaljuje in ponavlja, dokler se vse dimenzije na vzorcu ne ujemajo z dimenzijskimi na populaciji (Brick et al. 2003: 3–4). V empiričnem delu bom predstavila postopek rakinga, kjer eno dimenzijo predstavlja populacijska tabela spol x starost x izobrazba, drugo dimenzijo pa populacijska tabela za regijo.

### **4.3 Osnovno uteževanje v ESS**

Kadar so podatki iz različnih meritev (in različnih držav) združeni v analizi, je potrebno upoštevati spreminjajoči se design in populacijske uteži (Ganninger 2007: 3). V vzorčnih podatkih ESS sta že v osnovi vključeni dve uteži, ki naj bi jih uporabili pri vsaki analizi (Norwegian Social Science Data Services 2008b):

1. Prva je vzorčna utež (design weight ali *dweight*). V mnogih vzorčnih okvirjih, ki so jih izbrale sodelujoče države, niso imeli vsi posamezniki, starejši od 15 let, enake možnosti za izbor in so tako nekateri segmenti podcenjeni ali precenjeni (npr. regije). Vzorčna utež popravi ta nabor različnih verjetnosti in tako napravi vzorec bolj reprezentativen za posameznike v ciljni populaciji. To utež je potrebno vključiti v vsako analizo (tudi za tako imenovane neutrežene vzorce). Belgija, Slovenija in Švedska imajo te uteži enake 1.
2. Druga je populacijska utež (Population Size weight ali *pweight*). To utež uporabimo, kadar pregledujemo podatke za dve ali več držav skupaj. Večina držav je imela vzorec podobne velikosti, kljub različnim velikostim držav, in ta utež to popravi. Populacijska utež je koristna takrat, ko bi radi raziskavo ESS primerjali z drugimi mednarodnimi raziskavami.

»Na splošno so bile uteži za vsako državo izračunane po naslednjem postopku:  
1.)  $w = 1/(PROB1 * ... * PROBk)$  je  $n \times 1$  vektor uteži; k je odvisen od števila stopenj vzorca. Vse uteži so bile na koncu popravljene tako, da je vsota končne uteži enaka velikosti realiziranega vzorca (n), se pravi  $\text{weights} = n * w / \text{sum}(w)$ « (Jowell in Malnar 2006).

#### **4.4 Uteževanje pri prvi in drugi meritvi**

Uteževanje v prvih dveh meritvah so izvajali Vehovar (2006) ter Vehovar in Zupanič (2006), po katerih povzemam to poglavje. Za uteževanje so uporabili tri spremenljivke: spol (gender – G), starost (age – A) in izobrazba (education – E). Za potrebe uteževanja so bile le-te združene v naslednje kategorije:

Spol: dve kategoriji (moški, ženski).

Starost: tri kategorije (15–34 let, 35–54 let, 55 let in starejši).

Izobrazba: *nižja sekundarna ali manj* (iz ESS-spremenljivke 'edulvl' zajema naslednje vrednosti po ISCED 1997: nedokončana primarna izobrazba, primarna ali prva stopnja osnovne izobrazbe, nižja sekundarna, druga stopnja osnovne izobrazbe); *višja sekundarna* (dokončana višja sekundarna); *post sekundarna* (več kot 12 let, post sekundarna, a ne terciarna, terciarna izobrazba).

Za takšno starostno razdelitev so se odločili, ker na tak način dobimo celice, ki zadostujejo potrebni frekvenci 30 oziroma minimalno 10 enot. Poleg tega se je takšno združevanje izkazalo za najbolj učinkovito, saj so uteževane spremenljivke tako najbolj raznolike. Na podlagi danih populacijskih podatkov so uteževali s poststratifikacijsko metodo, kjer so imeli na razpolago tridimenzionalno tabelo (spol x starost x izobrazba), kjer pa le-ta ni bila na voljo, so uteževali z metodo rakinga.

Za vsako od kategorij uteževanih spremenljivk so bile izračunane uteži glede na razliko med strukturo populacije in strukturo vzorca. Po uteževanju so na vzorcu in izbranih spremenljivkah izračunali tudi hi-kvadrat in njegovo statistiko, VIF-faktor, pristranskost zaradi odgovorov ipd.

## 5 Podatkovni viri

Ključni problem, ki nastane pred začetkom uteževanja, je, kje lahko pridobimo kontrolne podatke, primerne za naš vzorec. Za uteževanje v Evropski družboslovni raziskavi potrebujemo vrednosti spremenljivk za celotno ciljno populacijo, da lahko primerno utežimo podatke.

Za meritev 1 (R1) in meritev 2 (R2) so bile uteži sprva pridobljene na podlagi ESS-jeve priloge Appendix 1: Population statistics and other documentation s pomočjo poststratifikacije, kjer so bile na voljo tridimenzionalne tabele za populacijske podatke, in z metodo rakinga tam, kjer teh tabel ni bilo na voljo. Je pa prihajalo do zapletov, saj so bili populacijski podatki ponekod nepopolni (z manjkajočimi podatki) in je bilo potrebno kontaktirati nacionalne koordinatorje za bolj popolne podatke (Vehovar 2006; Vehovar in Zupanič 2007). Poleg tega ESS svoje priloge dopolnjuje in osvežuje, kar lahko pripelje do nezmožnosti točne ponovitve uteževanja. Populacijski podatki iz ESS-jeve priloge torej niso vedno popolni, zato je potrebno preveriti še kakšno alternativno možnost za pridobitev populacijskih podatkov.

Poznamo več podatkovnih virov, kjer lahko pridobimo populacijske podatke, ki jih nato uporabimo pri uteževanju. Najboljše pa je uporabiti podatke iz enega samega vira, saj je »združevanje podatkov iz različnih virov lahko problematično, če se kvaliteta podatkov razlikuje. Herrera in Kapur pravita, da lahko kombinacija podatkov iz virov, ki imajo različne stopnje merske napake, vodi k pristranskim in nekonsistentnim rezultatom, če je napaka na spremenljivki, ki korelira z interesnimi spremenljivkami ali pa celo z odvisno spremenljivko« (Rydland et al. 2007: 5).

*Veljavnost in zanesljivost kontekstualnih spremenljivk sta pomembni tudi, če so ti podatki vključeni v raziskavo samo zato, da pridobimo širši družbeni kontekst. Vendar naj bi uporabniki podatkov ESS pazljivo preverili kvaliteto podatkov. Upoštevati je potrebno negotovost podatkov, morda izračunati merske napake,*

*stopnjo zaupanja. Nato je potrebno najti podatke najboljše kvalitete, če le-teh ni na voljo, morda pridobiti lastne podatke. In nazadnje kritično presoditi, kdaj so primerjalni podatki res primerljivi* (Ryland et al. 2007: 8).

Za uteževanje potrebujemo vrednosti kontrolnih spremenljivk, za katere Rydland et al. (2007) uporabljajo tudi izraz kontekstualni podatki (*angl. contextual data*). Le-ti naj bi opisovali populacijo, njene družbene in ekonomske značilnosti geografskih območij, skratka naj bi opisovali okolje, v katerem živijo posamezniki. V tej nalogi se osredotočam na štiri dimenzijske spol, starost, izobrazba in regija, ki po mojem mnenju najbolje klasificirajo populacijo. Poleg njih se sem šteje še dohodek, verska pripadnost itd., vendar je podatke za takšna vprašanja še težje dobiti.

Sledi predstavitev podatkovnih virov, ki so uporabljeni v analitičnem delu te diplomske naloge (ESS in Eurostat). Seveda obstajajo tudi drugi viri, ki pa niso tako primerni zaradi pomanjkljivih podatkov. Takšna je na primer Organizacija za ekonomsko sodelovanje in razvoj (Organisation for Economic Cooperation and Development – OECD), ki vsebuje mnogo podatkov, visoko ustreznih za ESS (na primer izobrazbeni sistem in demografija), vendar pa so podatki objavljeni le za države v tej organizaciji, kar izključuje mnoge države ESS (Rydland et al. 2007: 210). To pomanjkanje geografske pokritosti je glavni razlog, zakaj teh podatkov nisem raziskovala. Manjkajo jim namreč podatki za osem držav, ki sodelujejo v ESS, med njimi je tudi Slovenija.

Ker pa imamo lahko manjkajoče vrednosti tako na populaciji kot tudi na vzorcu, je potrebno oba vira podatkov (populacija in vzorec) prilagoditi, tako da sta ustrezena. O tem bom pisala v empiričnem delu svoje naloge.

## **5.1 ESS-dokumentacija**

Spletna stran ESS DATA vsebuje informacije o družbeno-kulturnem kontekstu znotraj vsake države. Glede na podatke, pridobljene na nacionalnih ravneh, so na strani prikazane populacijske statistike o spolu, starosti, izobrazbi in stopnji urbanizacije, poleg tega pa še specifične statistike ozadja, ki pomagajo vzpostaviti sovisnost modulov, kot npr. rasno sestavo populacije ali pa stopnje priseljevanja (Jowell et al. 2007: 25–26).

Na strani arhiva ESS-podatkov lahko pod zavihkom dokumentacija raziskave (*angl. Survey documentation*) najdemo informacije, povezane s potekom raziskave. Za vsako meritve najdemo tudi Prilogo A1: *populacijske statistike in druga dokumentacija* (ESS DATA 2003d, 2005c in 2007c), ki vsebuje širok spekter populacijskih statistik, ki so jih priskrbeli nacionalni koordinatorji sodelujočih držav. V tej datoteki so na voljo tudi podatki za usmerjanje analize anketnih podatkov in za metodološke namene (kot je na primer uteževanje).

Pri prehodu iz meritve 1 na meritve 2 so se koordinatorji raje kot za razvijanje nove podatkovne baze odločili za prenovitev in dopolnitev inventarja javnih elektronsko dostopnih informacij o kontekstu. Razpoložljiva podatkovna baza se razlikuje glede na zanesljivost, obsežnost ažurnih informacij, dostopnost itd., razmere glede kontekstualnih podatkov pa so se izboljšale glede na meritve 1. Naraščajoče število organizacij in institucij je priskrbelo profile držav, poleg tega pa mednarodne organizacije, npr. Eurostat, čedalje bolj implementirajo politiko omogočanja dostopa do podatkov brez stroškov (Jowell et al. 2006: 50-51).

Naloga nacionalnih ekip je bila zagotoviti nekaj ključnih statistik za pojasnjitev vprašalnikov, med katerimi so bile najbolj pomembne naslednje: starostna in spolna struktura, izobrazbena sestava ter stopnja urbanizacije (Kolsrud et al. 2007: 149). Vendar pa ni strogih pravil o kvaliteti in vrsti podatkov, ki naj bi jih nacionalni koordinatorji priskrbeli. Za vse meritve obstaja sicer dokument »Appendix 2, Classifications and coding standards« (klasifikacije in kodirni

standardi), kjer najdemo kodirni okvir za izobrazbo, vendar pa ni določeno, na kakšen način naj bodo ti podatki predstavljeni (kakšne tabele ipd.). To je lahko tudi razlog, da populacijske statistike iz ESS-priloge A1 med sabo pogosto niso primerljive in imajo več manjkajočih podatkov. Prav tako niso primerljive med državami, saj vsak nacionalni koordinator priskrbi podatke za svojo državo brez obzira na druge države. Prav tako z leti nastanejo razlike med državami, saj se raziskava spreminja, menjajo se agencije, ki izvajajo ankete za posamezno državo, lahko se zamenja nacionalni koordinator, ki je odgovoren za podatke, morda se spremeni celo način pristopa. So pa tudi problemi s podatki iz cenzusov, ki niso vedno standardizirani in so lahko celo zastareli.

## 5.2 Eurostat: Labour Force Survey (LFS)

Eurostat (Eurostat 2008c) je statistična pisarna evropske skupnosti, vzpostavljena 1953. Njena misija je zbrati in analizirati številke, ki jih pridobijo z različnih evropskih statističnih uradov, tako da lahko oskrbijo evropske institucije s primerljivi in skladnimi podatki, na podlagi katerih lahko potem definirajo, implementirajo in analizirajo politike skupnosti. Eurostat pokriva Evropsko unijo, njene članice in partnerje, podatki pa so objavljeni za različne teme in zbirke. Eurostatovo primarno delo je priprava natančnih, zanesljivih in ažurnih statističnih podatkov na evropskem nivoju (Eurostat 2007d: 20).

Eurostat objavlja širok razpon statističnih podatkov. Večji del jih pridobijo iz uradnih registrov, administrativnih zapisov, državnih poročil in cenzusov. Nekatere statistike temeljijo na raziskovah kot npr. Državna raziskava o zdravju (*angl. National health survey*) in Anketa o delovni sili (*angl. Labour force survey*) (Rydland et al. 2007: 158).

Anketa o delovni sili Evropske unije (*angl. Labour Force Survey*), v nadaljevanju tudi LFS, je četrtnetna raziskava z velikimi vzorci, ki pokriva vso populacijo v zasebnih gospodinjstvih v EU, državah EFTA razen Lichtensteina ter državah kandidatkah za EU. Dobavlja četrtnetne rezultate o delovno aktivnih prebivalcih,

starejših od 15 let, prav tako pa tudi za osebe, ki niso delovno aktivne. Velikost vzorca je okoli 1700 tisoč posameznikov (Eurostat 2007a).

Kar se tiče geografske pokritosti, LFS pokriva vsa ozemlja članic EU, države EFTA (brez Lichtensteina) ter Bolgarijo, Hrvaško in Romunijo. Obstajajo tudi odstopanja od ciljne populacije, in sicer se podatki za Ciper nanašajo samo na ozemlje, ki ga pokriva vlada ciprske republike, podatki za Francijo pa ne vključujejo provinc (Eurostat 2007c: 3).

Podatki iz Eurostata slovijo po visoki kvaliteti, sporne zadeve, ki bi lahko vplivale na kvaliteto, so večinoma dokumentirane. Eurostat je verjetno tudi najbolj pomemben vir združenih (aggregiranih) statistik za ESS-uporabnike, saj objavljajo visoko kakovostne podatke za širok spekter tem. Še dodatno, Eurostat je edina mednarodna statistična agencija, ki objavlja podatke na regionalnem nivoju. Večja slabost pa je slabša razpoložljivost podatkov za srednje in vzhodno evropske države (Rydland et al. 2007: 163).

## **5.3 Spremenljivke**

Sledi predstavitev spremenljivk in njihove vrednosti glede na ESS in LFS. Pri vsaki spremenljivki je najprej uvod, nato pa razлага situacije v ESS- in LFS-podatkih.

### **5.3.1 Spol**

Pri tej spremenljivki se pojavi najmanj težav. Vedno ima dve kategoriji, moški in ženski spol, občasno pa se pojavijo tudi manjkajoče vrednosti. V ESS-jevih populacijskih podatkih se zgodi, da imamo na razpolago dve tabeli, spol/starost in spol/izobrazba, z različnima številoma vseh oseb v obeh tabelah. Razlika med njima je manjkajoča vrednost za spol.

### **5.3.2 Starost**

Starost je v ESS-vzorcu izračunana iz dneva opravljenega intervjuja in letnice rojstva. Ciljna populacija raziskave ESS je prebivalstvo starejše od 15 let. Kot že omenjeno, vse države iz različnih razlogov ne morejo priskrbeti populacijskih podatkov za ciljno populacijo, lahko imajo dano neko zgornjo mejo. Velika Britanija, na primer, v vzorec vključuje samo prebivalce, stare do 65 let.

Eurostatove definicije (2007c) se nanašajo na ljudi, starejše od 15 let. Anketa o delovni sili (LFS) pokriva celotno populacijo, ki prebiva v državah članicah, z izjemo oseb, ki živijo v kolektivnih gospodinjstvih ali institucijah. Demografski podatki so zbrani za vse starostne skupine, odgovori na vprašanja, ki se nanašajo na delovni trg, pa se omejujejo na osebe nad 15 let. Za države EFTA (Islandija, Norveška in Švica) so populacijski podatki zbrani samo za starostne skupine znotraj delovnega trga, ki obsega osebe, stare od 15 do 75 let.

### **5.3.3 Izobrazba**

Izobrazbeni sistemi se od države do države razlikujejo. Vendar pa obstaja mednarodna lestvica, ki določa, kakšna vrsta/stopnja izobrazbe je potrebna, da se doseže določen izobrazbeni nivo. Ta lestvica se imenuje Mednarodna standardna klasifikacija izobraževanja (angl. *International Standard Classification of Education*) ISCED, ki jo vzdržuje UNESCO. Ustanovljena je bila v 70-ih letih 20. stoletja, da bi se olajšala statistična primerjava izobrazbe med različnimi državami. Leta 1997 so jo prenovili, tako da sedaj predstavlja mednarodni standard za vse organizirane in vzdrževane izobrazbene aktivnosti (Uis v Rydland et al. 2007: 59).

Smith et al. omenjajo (v Ryland et al. 2007: 59-64), da čeprav so mednarodni standardi jasni, je včasih njihovo integriranje v državne sisteme za države problematično. Države se razlikujejo glede na to, kaj definirajo kot izobrazbo oziroma kje so njene meje, kar lahko privede do problema s primerljivostjo. Meje izobrazbe lahko vključujejo tudi predšolsko izobrazbo in skrbstva, posebno

izobrazbo, poklicno izobrazbo ipd. Obstajajo tudi države, ki ne sledijo ISCED-standardom, lahko se spremenijo definicije izobrazbenih nivojev itd.

ESS v prilogi 3 (ESS DATA 2003e: 69) predstavi nabor vrednosti za izobrazbo, ki je malce pritejen ISCED-97: »0 Not completed primary education, 1 Primary or first stage of basic, 2 Lower secondary or second stage of basic, 3 Upper secondary, 4 Post secondary, non-tertiary, 5 First stage of tertiary, 6 Second stage of tertiary, 7 Refusal, 8 Don't know, 9 No answer.«<sup>1</sup>

ESS-jev nabor vrednosti tako temelji na ISCED 1997, kodiran pa je neposredno iz specifičnih spremenljivk za posamezno državo. Vsaka država tako v vzorcu vsebuje dve spremenljivki, ena je enaka za vse (EDULVL) in druga specifična za državo (za Belgijo npr. EDULVLBE).

Vendar izobrazba ni neproblematična. Države so morale razviti specifična vprašanja, da so pridobile nove izobrazbene kvalifikacije, ki so jih potem lahko podali v okvir ISCED (Kolsrud et al. 2007: 146).

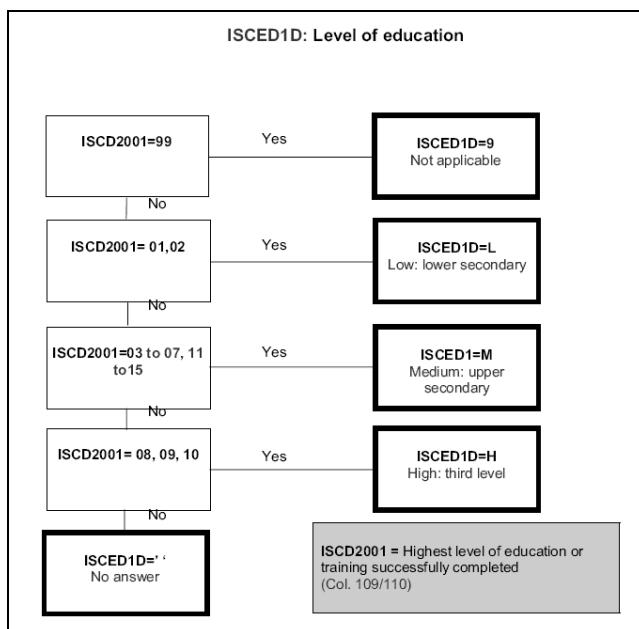
LFS za izobrazbo uporablja ISCED-klasifikacijo, se pa kodiranje spreminja glede na aktualno ISCED-klasifikacijo. Spremenljivke, ki jih uporabi LFS, so HATLEVEL in iz nje izpeljano HATLEV1D (LFS 2003). Vrednosti: »00 No education; 01 ISCED 1; 02 ISCED 2; 11 ISCED 3 (without distinction a, b or c possible); 03 ISCED 3c (shorter than 3 years); 04 ISCED 3c (3 years or more); 13 ISCED 3b; 14 ISCED 3a; 15 ISCED 4; 08 ISCED 5b; 09 ISCED 5a; 10 ISCED 6; 99 Not applicable (child less than 15 years); blank No answer« (LFS 2003: 30).

---

<sup>1</sup> V Sloveniji se je vprašanje za izobrazbo glasilo takole (ESS DATA 2003c: 27):

»**F6: Katera je zadnja šola, ki ste jo končali, redno ali izredno?** 01 – nedokončana osnovna šola; 02 – dokončana osnovna šola; 03 – 2- ali 3- letna poklicna šola; 04 – splošna gimnazija, poklicna gimnazija, štiriletna strokovna šola; 05 – 2-letna višja (strokovna) šola; 06 – visoka šola ali fakulteta; 07 – magisterij, doktorat; 88 - ne vem.«

HATLEV1D je izpeljana v tri skupine, nizka, srednja in visoka izobrazba, kot vidimo na Slika 5.3.3.1:



**Slika 5.3.3.1: Trije izobrazbeni nivoji v LFS**

Vir: LFS 2003: 35.

Pri uteževanju v meritvi 1 in meritvi 2 so za izobrazbo uporabili prav tako tri izobrazbene skupine, vendar so se razlikovale od LFS-jeve izpeljane spremenljivke, kategorija '15 ISCED 4' je bila vključena v tretjo skupino (višja izobrazba), LFS pa jo vključuje v drugo skupino. V tej diplomski nalogi sem izobrazbo združevala na LFS-način, četrta vrednost spremenljivke EDULVL 'Post secondary, non-tertiary' je bila vključena v drugo skupino.

Tabela 5.3.3.1 predstavlja vrednosti specifičnih izobrazbenih spremenljivk za izbrane države ter zrduževanje v meritvah za LFS. V vsaki meritvi so vse tri države vsebovale spremenljivko EDULVL, na podlagi katere sem združevala v tri izobrazbene kategorije.

**Tabela 5.3.3.1: Izobrazbene strukture za BE, SI in SE**

Država	Ime sprem.	R1	R2	R3	LFS (sprem.. HATLEVEL)
Belgija	EDLVBE and EDULVL	01 Niet voltooid lager onderwijs 02 Lager onderwijs 03 Lager beroepsonderwijs 04 Lager algemeen secundair onderwijs 05 Hoger secundair onderwijs 06 Hoger secundair technisch, zevend beroepsopleiding 07 Hoger algemeen secundair 08 Hoger onderwijs korte type 09 Hoger onderwijs van het lange type 10 Universiteit 11 Doctoraal en postdoctoraal 12 Andere, 77 Refusal, 88 Don't know, 99 No answer	01 Niet voltooid lager onderwijs 02 Lager onderwijs 03 Lager beroepsonderwijs 04 Lager algemeen secundair onderwijs 05 Hoger secundair onderwijs 06 Hoger secundair technisch, zevend beroepsopleiding 07 Hoger algemeen secundair 08 Hoger onderwijs korte type 09 Hoger onderwijs van het lange type 10 Universiteit 11 Doctoraal en postdoctoraal 12 Andere, 77 Refusal, 88 Don't know, 99 No answer	EDULVL	R1, R2: L - 11, 21 M - 31, 32, 43 H - 51, 52, 60  R3: L - 0, 11, 21 M - 31, 32, 43 H - 51, 52, 60
Slovenija	EDLVSI and EDULVL	EDULVL	EDULVL	01 Not completed primary education 02 Primary or first stage of basic 03 Lower secondary or second stage of basic 04 Upper secondary 05 Post secondary, non-tertiary 06 First stage of tertiary 07 Second stage of tertiary 77 Refusal 88 Don't know 99 No answer	R1: L - 11, 21 M - 31, 32 H - 51, 52, 60  R2, R3: L - 0, 11, 21 M - 31, 32 H - 51, 52, 60
Švedska	EDLVSE EDLVSE EDLVASE and EDULVL	01 Not finished elementary school 02 Elementary school, old 03 Elementary school 04 Lower secondary and elementary school, old 05 Vocational school 1963-1970 06 2 year high school 07 3-4 year high school prior 1995 08 Vocational high school after 1992 09 Theoretical high school after 1992 10 University, no exam 11 University, exam less than 3 years 12 University, exam more than 3 years 77 Refusal 88 Don't know 99 No answer	01 Not finished elementary school 02 Elementary school, old 03 Elementary school 04 Lower secondary and elementary school, old 05 Vocational school 1963-1970 06 2 year high school 07 3-4 year high school prior 1995 08 Vocational high school after 1992 09 Theoretical high school after 1992 10 University, no exam 11 University, exam less than 3 years 12 University, exam more than 3 years 77 Refusal 88 Don't know 99 No answer	01 Ej avslutad folkskola/grundskola 02 Folkskola 03 Grundskola/enhetsskola 04 Realskola/Flickskola 05 Fackskola (1963-1970) 06 2-årigt gymnasium 07 3- eller 4 årigt gymnasium (före 1995) 08 Yrkessinriktat gymnasium (efter 1992) 09 Teoretiskt gymnasium (efter 1992) 10 Universitet/högskola ej examen 11 Universitet/högskola, mindre än 3 år 12 Universitet/högskola > 3år med examen 13 Forskarutbildning 66 Not applicable 77 Refusal, 88 Don't know, 99 No answer	R1: L - 11, 21 M - 36, 43 H - 51, 52, 60  R2, R3: L - 11, 21 M - 30, 43 H - 51, 52, 60

**Vir: ESS DATA 2003e, 2005d, 2007d; Eurostat 2007b; LFS 2003: 59-65.**

### 5.3.4 Regija

Zakaj regija? V prvih dveh meritvah ESS se niso odločili za uteževanje na regijo. V literaturi pa raziskovalci (Fowler 2002: 42) ugotavljajo, da so stopnje odgovorov (*angl. response rates*) v splošnem višje na podeželju kot v mestih, različne vrste dostopnosti pa pridobivajo pristranskost, ki je povezana z neodgovori. Tako Fowler pravi, da imajo nacionalne ankete z osebnimi intervjuji skoraj vedno nižjo stopnjo odgovorov v mestnih centrih kot pa na obrobju mest ali podeželju, in meni, da za to obstajajo trije razlogi. Prvič, stopnja »hard-to-find« posameznikov je višja v mestnih centrih. Drugič, naraščajoča skupina posameznikov živi v stanovanjskih stavbah, do koder imajo anketarji otežen direkten dostop. In tretjič, v mestu je več območij, kjer so nočni obiski neprijetni za anketarje in nimajo tako dobre možnosti najti ljudi, ki jih drugače teže najdejo na domu (Fowler 2002: 42–43). Uteževanje na regijo torej lahko zmanjša pristranskost.

Podatke o regijah lahko najdemo na različne načine. Lahko se kontaktira nacionalne koordinatorje, se jih izbrska na internetnih straneh mednarodnih organizacij ali statističnih agencij itd. Ob iskanju je potrebno upoštevati, da ima vsaka država različne regije. Prav tako pa obstaja tudi tu mednarodni standard. Kot piše na spletni strani EU (2008), je klasifikacijo statističnih teritorialnih enot v Evropski uniji s kratico NUTS<sup>2</sup> ustanovil Eurostat pred več kot 30 leti, da bi zagotovili enotno razdelitev teritorialnih enot za razvoj regionalnih statistik Evropske unije. Klasifikacija NUTS je hierarhično razdeljena na tri ravni: NUTS 1, NUTS 2 in NUTS 3 (glej tudi Statistični urad 2008). Tako se ozemlje držav deli na prvi ravni na enote NUTS 1, te se nadalje členijo na enote ravni NUTS 2, tretja raven pa so enote NUTS 3.

Države morajo pri razdelitvi svojega ozemlja na enote NUTS upoštevati normativna merila (število prebivalstva), ki so določena v uredbi o NUTS. To pomeni, da ni nujno, da vsaka država uporabi vse tri ravni hierarhične

---

<sup>2</sup> Francosko *nomenclature des unites territoriales statistiques*.

razdelitve. Luksemburg je npr. država, katere celotno ozemlje predstavlja tako enoto na ravni NUTS 1 kot tudi na ravneh NUTS 2 in NUTS 3. Povprečna velikost enot NUTS je opredeljena z naslednjimi pragi:

**Tabela 5.3.4.1: Ravni NUTS in število prebivalcev**

Raven	Najmanj	Največ
NUTS 1	3 milijone	7 milijonov
NUTS 2	800.000	3 milijone
NUTS 3	150.000	800.000

Vir: Statistični urad 2008

Obstajajo vsaj trije problemi, ki so povezani z uporabo NUTS v primerjalni analizi. 1) Čeprav naj bi ta klasifikacija formirala usklajen sistem regionalnih enot, še vedno ostaja zelo heterogena. Tako lahko znotraj države velikosti znotraj ene ravni močno variirajo, prav tako se lahko razlikujejo v drugih pogledih. 2) Ponekod so meje med teritorialnimi enotami začrtane tako, da v resnici podajajo zavajajoče regionalne podatke. 3) NUTS struktura se je spremenjala skozi leta. Spremembe v državah lahko vodijo v spremembe v NUTS klasifikaciji (Rydland et al. 2007: 20–22).

Najbolj pogosta vrsta podatkov, ki jih uporabimo za merjenje konteksta so ponavadi regionalni podatki. Le-ti opišejo značilnosti ozemlja in prostorskih enot. Že od začetka ESS je načrtovano, da se kontekstualne informacije pridobi od vsake sodelujoče države, vendar je akademsko vodene regionalne in kontekstualne podatkovne baze težko dobiti. Eurostat, kot ena izmed uradnih statističnih agencij, je zbral vse večje podatkovne baze na regionalnem nivoju (NUTS), ki pa, ne glede na njihovo dragoceno vrednost, niso dovolj za potrebe akademske skupnosti, delno zaradi njihovega konteksta, delno zaradi njihovega okvira (niso vključene vse evropske države) in delno zaradi njihove omejene uporabnosti (Stoop 2007: 19).

V ESS ima vsaka država svojo spremenljivko za regijo. Večinoma sicer države sledijo eni izmed NUTS ravni, vendar obstaja tudi nekaj držav s specifičnimi regionalnimi razdelitvami. LFS vsebuje podatke samo za raven NUTS 2. Tabela 5.3.4.2 predstavlja primerjavo za tri proučevane države.

**Tabela 5.3.4.2: Primerjava regionalnih kategorij za vse tri meritve**

Država	Ime sprem.	R1	R2	R3	NUTS 2 (LFS)	Komentar
Belgia	REGIONBE	01 Flemish region 02 Brussels region 03 Walloon region 99 Not available	1 Flemish region 2 Brussels region 3 Walloon region 999 Not available	1 Flemish region 2 Brussels region 3 Walloon region 999 Not available	10 Région de Bruxelles-Capitale 21 Prov. Antwerpen 22 Prov. Limburg 23 Prov. Oost-Vlaanderen 24 Prov. Vlaams Brabant 25 Prov. West-Vlaanderen 31 Prov. Brabant Wallon 32 Prov. Hainaut 33 Prov. Liège 34 Prov. Luxembourg 35 Prov. Namur	ESS ima regije na ravni NUTS 1, LFS pa na ravni NUTS 2.  Lahko združimo NUTS 2 ravni na NUTS 1: 21, 22, 23, 24, 25 = 1 Flemish region 10 = 2 Brussels region 31, 32, 33, 34, 35 = 3 Walloon region
Slovenija	REGIONS1	01 Gorenjska 02 Goriska 03 Jugovzhodna Slovenija 04 Koroska 05 Notranjsko-kraska 06 Obalno-kraska 07 Osrednjeslovenska 08 Podravska 09 Pomurska 10 Savinjska 11 Spodnjeposavska 12 Zasavska 99 Not available	01 Gorenjska 02 Goriska 03 Jugovzhodna Slovenija 04 Koroska 05 Notranjsko-kraska 06 Obalno-kraska 07 Osrednjeslovenska 08 Podravska 09 Pomurska 10 Savinjska 11 Spodnjeposavska 12 Zasavska 999 Not available	01 Gorenjska 02 Goriska 03 Jugovzhodna Slovenija 04 Koroska 05 Notranjsko-kraska 06 Obalno-kraska 07 Osrednjeslovenska 08 Podravska 09 Pomurska 10 Savinjska 11 Spodnjeposavska 12 Zasavska 999 Not available	01 Vzhodna Slovenija 02 Zahodna Slovenija	ESS ima regije na NUTS 3 ravni, LFS pa na NUTS 2 ravni.  Lahko združimo NUTS 3 v NUTS 2: 03, 04, 05, 08, 09, 10, 11, 12 = 01 Vzhodna Slovenija 01, 02, 06, 07 = 02 Zahodna Slovenija
Švedska	REGIONSE	01 Stockholm 02 Östra Mellansverige 03 Sydsverige 04 Norra Mellansverige 05 Mellemta Norrland 06 Övre Norrland 07 Småland och Öarna 08 Västsverige 99 Not available	1 Stockholm 2 Östra Mellansverige 3 Sydsverige 4 Norra Mellansverige 5 Mellemta Norrland 6 Övre Norrland 7 Småland och Öarna 8 Västsverige 999 Not available	1 Stockholm 2 Östra Mellansverige 3 Sydsverige 4 Norra Mellansverige 5 Mellemta Norrland 6 Övre Norrland 7 Småland och Öarna 8 Västsverige 999 Not available	01 Stockholm 02 Östra Mellansverige 09 Småland med öarna 04 Sydsverige 0A Västsverige 06 Norra Mellansverige 07 Mellersta Norrland 08 Övre Norrland  03 Småland med öarna (till 1998) 05 Västsverige (till 1998)	Enaka regionalna raven, vendar imajo LFS 2002, 2004, 2006 podatki regijo 0, kar verjetno predstavlja manjkojočo vrednost.

Viri: ESS DATA 2003e, 2005d, 2007d; Eurostat 2008b; European Communities 2008.

## 5.4 Priprava podatkov

ESS-podatke sem pridobila s spletnne strani ESS (ESS DATA 2003a, 2003d, 2005a, 2005c, 2007a in 2007c).

LFS-podatke mi je posredoval mentor red. prof. dr. Vasja Vehovar, ki jih je v obliki csv preko elektronske pošte dobil z Eurostatovega urada v Luksemburgu (Eurostat 2008a). Podatki so bili za vse države za leta 2002, 2004 in 2006. V Excelu je sledila priprava podatkov s funkcijo »besedilo v stolpc«, razbitjem podatkov iz enega lista na več listov (glede na leta in države, ki so sodelovale v ESS-raziskavi), nazadnje pa še vrtilna tabela za spol/starost/izobrazbo za prebivalce v tisočih (*value*), ki napravi takšno tabelo. Za izobrazbo sem zaradi boljše primerljivosti vzela že agregirano spremenljivko HATLEV1D. Za primer bom predstavila Slovenijo v letu 2004, ki ima manjkajoče vrednosti za izobrazbo, obenem pa ima samo dve regiji (razlaga sledi v nadaljevanju).

**Tabela 5.3.4.1: Slovenija, LFS 2004**

Sum of VALUE		HATLEV1D				
SEX	AGE	1. Low	2. Medium	3. High	No answer (blank)	Grand Total
1.Males	15–19	54.3151	10.8344		0.5521	65.7017
	20–24	10.6896	64.5297	1.1762		76.3955
	25–29	7.4553	57.2839	11.8519		76.5910
	30–34	10.2061	52.3934	13.6675		76.2670
	35–39	11.7578	52.7919	9.9134		74.4631
	40–44	11.7491	48.7912	9.9572		70.4975
	45–49	18.6034	56.5180	14.4021		89.5236
	50–54	17.0277	45.5969	14.2959		76.9205
	55–59	12.1558	34.8951	9.3932		56.4441
	60–64	11.2316	31.7286	8.6210		51.5812
	65+	35.1727	59.7307	18.7156		113.6190
1.Males Total		200.3642	515.0938	111.9940	0.5521	828.0041
2.Females	15–19	54.4003	8.7628		0.3487	63.5117
	20–24	4.5587	64.5654	3.0412		72.1652
	25–29	3.8831	43.9815	24.0739		71.9384
	30–34	7.7112	41.2405	22.5517		71.5034
	35–39	11.1730	48.7047	14.6641		74.5417
	40–44	14.6178	39.6271	15.7984		70.0433
	45–49	26.0180	42.9412	14.4805		83.4396
	50–54	22.4414	37.7330	13.1228		73.2972
	55–59	21.5158	23.9884	9.7309		55.2351
	60–64	26.3559	25.5877	6.1587		58.1023
	65+	116.0859	57.5955	10.9191		184.6005
2.Females Total		308.7609	434.7278	134.5412	0.3487	878.3786
(blank)	(blank)					
(blank) Total						
Grand Total		509.1251	949.8216	246.5352	0.9008	1706.3827

Da bi pridobila celoten N in primerno tabelo za uteževanje, sem množila celotno tabelo s 1000 ter starost združila v primerne kategorije.

**Tabela 5.3.4.2: Združeni LFS-podatki: Slovenija 2004 – spol/starost/izobrazba**

spol	starost	izobrazba				skupaj
		nizka	srednja	visoka	manjkajoča	
moški	15–34	82666.1	185041.4	26695.6	552.1	294955.2
	35–54	59138.0	203698.1	48568.6	0.0	311404.7
	55+	58560.1	126354.3	36729.8	0.0	221644.2
	skupaj	200364.2	515093.8	111994.0	552.1	828004.1
ženske	15–34	70553.2	158550.1	49666.8	348.7	279118.8
	35–54	74250.1	169006.0	58065.8	0.0	301321.9
	55+	163957.6	107171.6	26808.7	0.0	297937.9
	skupaj	308760.9	434727.8	134541.2	348.7	878378.6
	skupaj	509125.1	949821.6	246535.2	900.8	1706382.7

Da bi preverila, ali je tabela res primerna za uteževanje, sem najprej preverila, kakšno število enot bi dobila v namišljenem vzorcu z 2000 enotami (vsi ESS-vzorci se namreč gibljejo okoli tega števila), saj velja, da

*so npr. ocene deležev, ki se nanašajo na podskupine z manj kot  $n_a=10$  elementi, povsem neprimerne za javno objavo. Pri tem je nepomembno, kako velika je populacija in kako velik je vzorec, če je le delež relativno majhen. .../ Meja  $n_a = 30$  elementov – včasih se navaja tudi  $n_a = 20$  elementov – je pogosto navedena kot najmanjša velikost podskupine, ki sploh lahko nastopa kot samostojna enota v statističnih analizah (Little in Rubin 1987). Pri tem ne gre le za analizirane podskupine v vzorcu, ampak tudi za celice, v katerih opravimo uteževanje ali druge korekcijske metode (Kalton in Vehovar 2001: 135–136).*

**Tabela 5.3.4.3: Tabela 5.3.4.2 projicirana na 2000 enot (namišljeni vzorec)**

spol	starost	izobrazba				skupaj
		nizka	srednja	visoka	manjkajoča	
moški	15–34	96.89	216.88	31.29	0.65	345.71
	35–54	69.31	238.75	56.93	0.00	364.99
	55+	68.64	148.10	43.05	0.00	259.78
	skupaj	234.84	603.73	131.26	0.65	970.48
ženske	15–34	82.69	185.83	58.21	0.41	327.15
	35–54	87.03	198.09	68.06	0.00	353.17
	55+	192.17	125.61	31.42	0.00	349.20
	skupaj	361.89	509.53	157.69	0.41	1029.52
	skupaj	596.73	1113.26	288.96	1.06	2000.00

Vidim lahko, da z izjemo v stolpcu z manjkajočimi vrednosti (sivi celici) nikjer ni celice z manj kot 30 enotami. Populacijska tabela spol/starost/izobrazba je tako primerna za uteževanje.

Kasneje sem dodala še podatke za regijo in pripravila tabelo regija/starost/spol/izobrazba<sup>3</sup>. Regijski podatki so na voljo samo v obliki NUTS 2 za vse države in so označeni s številkami. Preden se začne z uteževanjem, je potrebno še pogledati, katero regijo predstavlja kakšna številka. Takšne tabele za vse tri države in vsa tri leta so v prilogi, tu pa, kot je bilo že omenjeno, predstavljam Slovenijo v letu 2004, ker ima od treh izbranih držav najmanj regij.

**Tabela 5.3.4.4: Slovenija, LFS 2004**

Sum of VALUE		SEX HATLEV1D				1.Males				2.Females				2.Females			(blank)	(blank)	Grand Total
region	AGE	1. Low	2. Medium	3. High	No answer	Total	1. Low	2. Medium	3. High	No answer	Total	(blank)	(blank)	Total	(blank)	(blank)			
1	15-19	29.5552	6.0145		0.1178	35.6875	29.2132	5.5588		0.2400	35.0119						70.6994		
	20-24	5.0536	35.9437	1.1028		42.1001	2.5390	33.9686	1.8874		38.3951						80.4951		
	25-29	3.6606	31.2484	5.9674		40.8764	2.0398	25.8683	12.4917		40.3997						81.2761		
	30-34	4.7230	30.8711	6.2465		41.8405	4.2609	19.6289	12.0221		35.9119						77.7524		
	35-39	6.8242	28.8310	3.7827		39.4379	6.8849	27.3037	7.6425		41.8311						81.2690		
	40-44	8.4280	25.9739	3.8138		38.2156	9.6648	22.3709	7.6225		39.6582						77.8739		
	45-49	11.9260	31.7981	7.3394		51.0635	15.0455	22.1235	6.7982		43.9672						95.0307		
	50-54	11.6644	24.8968	6.0552		42.6164	14.8228	18.3736	5.3904		38.5867						81.2031		
	55-59	7.3606	18.1570	4.2646		29.7821	13.0243	12.0632	4.0324		29.1199						58.9020		
	60-64	7.8577	17.5414	3.8902		29.2893	17.2004	13.5047	2.6453		33.3504						62.6397		
	65+	22.0157	30.1044	6.3916		58.5117	66.3442	26.8494	4.1485		97.3421						155.8538		
1 Total		119.0689	281.3802	48.8540	0.1178	449.4210	181.0399	227.6134	64.6809	0.2400	473.5741						922.9951		
2	15-19	24.7599	4.8200		0.4343	30.0142	25.1871	3.2040		0.1087	28.4998						58.5140		
	20-24	5.6360	28.5861	0.0734		34.2955	2.0196	30.5968	1.1538		33.7702						68.0656		
	25-29	3.7947	26.0355	5.8845		35.7147	1.8433	18.1132	11.5822		31.5387						67.2534		
	30-34	5.4831	21.5223	7.4211		34.4265	3.4503	21.6116	10.5297		35.5915						70.0180		
	35-39	4.9336	23.9610	6.1307		35.0252	4.2880	21.4010	7.0216		32.7106						67.7358		
	40-44	3.3211	22.8173	6.1434		32.2819	4.9530	17.2562	8.1759		30.3851						62.6670		
	45-49	6.6774	24.7199	7.0628		38.4600	10.9725	20.8177	7.6823		39.4725						77.9325		
	50-54	5.3633	20.7001	8.2407		34.3041	7.6186	19.3595	7.7325		34.7105						69.0147		
	55-59	4.7953	16.7381	5.1285		26.6619	8.4915	11.9253	5.6985		26.1153						52.7772		
	60-64	3.3739	14.1872	4.7308		22.2920	9.1555	12.0830	3.5134		24.7519						47.0439		
	65+	13.1570	29.6262	12.3240		55.1073	49.7417	30.7461	6.7706		87.2584						142.3656		
2 Total		81.2953	233.7136	63.1399	0.4343	378.5831	127.7211	207.1144	69.8603	0.1087	404.8044						783.3876		
(blank)																			
(blank) Total																			
Grand Total		200.3642	515.0938	111.9940	0.5521	828.0041	308.7609	434.7278	134.5412	0.3487	878.3786						1706.3827		

Ponovim postopek za pripravo tabele za uteževanje in preverim na namišljenem vzorcu 2000 enot.

<sup>3</sup> Drugačen vrstni red spremenljivk je zaradi boljše preglednosti.

**Tabela 5.3.4.5: Združeni LFS-podatki: Slovenija 2004 – regija/starost/spol/izobrazba**

N		spol izobrazba				Skupaj				Skupna vsota			
regija	starost	moški		ženske		moški	ženske	skupaj moški	skupaj ženske				
		nizka	srednja	visoka	manjkajoča								
1	15–34	42992.4	104077.7	13316.7	117.8	160504.6	38052.9	85024.6	26401.2	240.0	149718.7		
	35–54	38842.6	111499.8	20991.1	0.0	171333.5	46418.0	90171.7	27453.6	0.0	164043.3		
	55+	37234.0	65802.8	14546.4	0.0	117583.2	96568.9	52417.3	10826.2	0.0	159812.4		
1 skupaj		119069.0	281380.3	48854.2	117.8	449421.3	181039.8	227613.6	64681.0	240.0	473574.4		
2	15–34	39673.7	80963.9	13379.0	434.3	134450.9	32500.3	73525.6	23265.7	108.7	129400.3		
	35–54	20295.4	92198.3	27577.6	0.0	140071.3	27832.1	78834.4	30612.3	0.0	137278.8		
	55+	21326.2	60551.5	22183.3	0.0	104061.0	67388.7	54754.4	15982.5	0.0	138125.6		
2 skupaj		81295.3	233713.7	63139.9	434.3	378583.2	127721.1	207114.4	69860.5	108.7	404804.7		
Skupna vsota		200364.3	515094.0	111994.1	552.1	828004.5	308760.9	434728.0	134541.5	348.7	878379.1		
											1706383.6		

**Tabela 5.3.4.6: Tabela 5.3.4.5 projicirana na 2000 enot (izmišljeni vzorec)**

n		spol izobrazba				Skupaj				Skupna vsota			
regija	starost	moški		ženske		moški	ženske	skupaj moški	skupaj ženske				
		nizka	srednja	visoka	manjkajoča								
1	15–34	50.39	121.99	15.61	0.14	188.12	44.60	99.65	30.94	0.28	175.48		
	35–54	45.53	130.69	24.60	0.00	200.81	54.41	105.69	32.18	0.00	192.27		
	55+	43.64	77.13	17.05	0.00	137.82	113.19	61.44	12.69	0.00	187.31		
1 skupaj		139.56	329.80	57.26	0.14	526.75	212.19	266.78	75.81	0.28	555.06		
2	15–34	46.50	94.90	15.68	0.51	157.59	38.09	86.18	27.27	0.13	151.67		
	35–54	23.79	108.06	32.32	0.00	164.17	32.62	92.40	35.88	0.00	160.90		
	55+	25.00	70.97	26.00	0.00	121.97	78.98	64.18	18.73	0.00	161.89		
2 skupaj		95.28	273.93	74.00	0.51	443.73	149.70	242.75	81.88	0.13	474.46		
Skupna vsota		234.84	603.73	131.26	0.65	970.48	361.89	509.53	157.69	0.41	1029.52		
											2000.00		

Še vedno ni nobene celice z manj kot 10 enotami (z izjemo manjkajočih vrednosti), a obstaja 5 celic z manj kot 20 enotami (temno sive) in 4 celice z manj kot 30 enotami (svetlo sive). Glede na to, da ima večina držav več kot dve regiji, takšna tabela ni primerna za uteževanje, saj bi bile vzorčne celice premajhne za analizo.

Za uteževanje glede na spol, starost, izobrazbo in regijo sem zato izbrala raking na poststratifikacijsko tabelo spol/starost/izobrazba in regijo. Preden se uteževanja lotimo, je potrebno preveriti še populacijske podatke in manjkajoče vrednosti, kar vsebuje tudi moji hipotezi. To bom preverjala v naslednjem poglavju, empirični analizi.

## 6 Empirična analiza

### 6.1 Populacijski podatki ESS vs. LFS

Najprej bom predstavila populacijske podatke, kot jih imamo na voljo v ESS-prilogi A1: populacijske statistike in druga dokumentacija. Te podatke bom zaradi preglednosti ohranila v jeziku, v katerem so navedeni v viru. Nato bom prikazala pretvorbo teh podatkov v obliko, ki je primerna za uteževanje ali raking. Tako bom predstavila podatke za vse tri meritve, priložila pa jim bom še podatke LFS v dveh tabelah, ustreznih za uteževanje: spol/starost/izobrazba in regija. Celotni podatki LFS v obliki 4D-tabele za vse spremenljivke so zaradi obsežnosti prikazani v prilogi B. Na koncu vsake države bodo izračunani odstotki za kategorije za spol, starost in izobrazbo, ki sem jih na podlagi prejšnjih virov (Vehovar 2006; Vehovar in Zupanič 2007) določila za uteževanje. Vse skupaj bom interpretirala in podala svojo ugotovitev glede tega, kateri vir je primernejši za uteževanje.

#### 6.1.1 Belgija

V prvi meritvi so belgijski nacionalni koordinatorji poslali podatke o razdelitvi prebivalcev po regijah glede na spol/starost ter glede na spol/izobrazbo.

Tabela 6.1.1.1: Populacijski podatki: Belgija R1 – spol/starost

Table 1 of age by gender			
Controlling for region=Brussels			
age	gender	Total	
Frequency	Female	Male	
15–24	60745	58706	119451
25–34	82572	84038	166610
35–44	71516	73920	145436
45–54	61746	58752	120498
55–64	47244	43059	90303
65–74	44713	33728	78441
75 pl	54348	25730	80078
Total	422884	377933	800817

Table 2 of age by gender			
Controlling for region=Flanders			
age	gender	Total	
Frequency	Female	Male	
15–24	350775	364767	715542
25–34	391137	402712	793849
35–44	466626	482445	949071
45–54	407503	420476	827979
55–64	329308	326153	655461
65–74	312005	270295	582300
75 pl	278329	164024	442353
Total	2535683	2430872	4966555

Table 3 of age by gender			
Controlling for region=Wallonia			
age	gender	Total	
Frequency	Female	Male	
15–24	201766	209317	411083
25–34	221155	225546	446701
35–44	253620	255132	508752
45–54	238712	238258	476970
55–64	170462	159920	330382
65–74	171234	135958	307192
75 pl	169349	86679	256028
Total	1426298	1310810	2737108

Vir: ESS DATA 2003d: 6-7.

**Tabela 6.1.1.2: Populacijski podatki: Belgija R1 spol/izobrazba**

Table 1 of educat by gender			Table 2 of educat by gender			Table 3 of educat by gender					
Controlling for region=Brussels			Controlling for region=Flanders			Controlling for region=Wallonia					
educat	gender	Total	educat	gender	Total	educat	gender	Total			
Frequency	Female	Male	Frequency	Female	Male	Frequency	Female	Male			
1_Primary	117634	85488	203122	1_Primary	722343	543622	1265965	1_Primary	417851	319539	737390
2_low_secund	88473	79143	167616	2_low_secund	527887	554805	1082692	2_low_secund	339958	333060	673018
3_High_secund	106604	101225	207829	3_High_secund	750224	797724	1547948	3_High_secund	381898	401991	783889
4_not_un_short	44763	27329	72092	4_not_un_short	351691	238882	590573	4_not_un_short	195203	118846	314049
5_not_un_long	12484	11821	24305	5_not_un_long	52203	88278	140481	5_not_un_long	20987	25497	46484
6_university	48739	67238	115977	6_university	119454	195235	314689	6_university	64832	104998	169830
Total	418697	372244	790941	Total	2523802	2418546	4942348	Total	1420729	1303931	2724660

Vir: ESS DATA 2003d: 5.

V tabelah so kategorije spremenljivk označene na naslednji način:

Spol: *male* – moški in *female* – ženske.

Starost: starostni razredi z razmakom desetih let, od 15 let do 75 in več.

Izobrazba: *1\_Primary* – nekončana primarna izobrazba ali primarna izobrazba,  
*2\_low\_second* – nižja sekundarna izobrazba,  
*3\_high\_second* – višja sekundarna izobrazba,  
*4\_not\_un\_short* – postsekundarna, ne terciarna izobrazba,  
*5\_not\_un\_long* – prva stopnja terciarne izobrazbe,  
*6\_university* – druga stopnja terciarne izobrazbe.

Podatke iz tabel sem združila v tri tabele, primerne za raking. Izobrazbo sem združila v tri razrede, ki se najbolj ujemajo s podatki LFS.

**Tabela 6.1.1.3: Združeni podatki iz ESS-priloge, primerni za raking (spol/izobrazba)**

izobrazba spol	nizka (1 + 2)	srednja (3)	visoka (4 + 5)	Skupaj
moški	1915657	1300940	878124	4094721
ženske	2214146	1238726	910356	4363228
Skupaj	4129803	2539666	1788480	8457949

**Tabela 6.1.1.4: Združeni podatki iz ESS, Belgija R1, primerni za raking (spol/starost)**

starost spol	15–34	35–54	55+	Skupaj
moški	1345086	1528983	1245546	4119615
ženske	1308150	1499723	1576992	4384865
Skupaj	2653236	3028706	2822538	8504480

Za regijo lahko sestavimo dve tabeli, odvisno od tega, katero izberemo za osnovno.

**Tabela 6.1.1.5: Združeni podatki iz ESS, Belgija R1, primerni za raking (regija)**

	spol/starost	ali		spol/izobrazba
Brussels	800817		Brussels	790941
Flanders	4966555		Flanders	4942348
Wallonia	2737108		Wallonia	2724660
Skupaj	8504480		Skupaj	8457949

Glede na populacijske podatke lahko izvedemo postopek rakinga na tri tabele. Vendar pa imamo, kot lahko vidimo v tabeli za regijo, v tabelah različno skupno vsoto. Če bi torej uporabili te populacijske podatke za uteževanje, bi bilo potrebno v tabelo za izobrazbo vstaviti manjkajoče vrednosti, tako da bi se skupni vsoti ujemali. Spremenljivke so drugače primerljive s spremenljivkami v vzorcu, tudi regija je na nivoju NUTS 1.

Združeni LFS-podatki na podlagi 4D-tabele za Belgijo 2002. Spol in izobrazba ostaneta nespremenjena, starost sem združila v tri razrede, regijo pa iz NUTS 2 združila v NUTS 1 po kategorijah, opisanih v Tabela 5.3.4.2. Podatki so primerni za raking na tabeli spol/starost/izobrazba in regija.

**Tabela 6.1.1.6: Združeni LFS-podatki: Belgija 2002 – spol/starost/izobrazba**

spol	starost	izobrazba			skupaj
		nizka	srednja	visoka	
moški	15–34	511999.3	555170.0	277919.0	1345088.2
	35–54	608267.9	500005.6	420708.8	1528982.3
	55+	791597.5	257070.0	196879.9	1245547.4
	skupaj	1911864.6	1312245.5	895507.7	4119617.9
ženske	15–34	420977.3	522655.1	364517.5	1308149.9
	35–54	604715.6	484903.5	410105.4	1499724.5
	55+	1159063.2	266296.4	151630.0	1576989.7
	skupaj	2184756.2	1273855.0	926252.9	4384864.1
skupaj		4096620.8	2586100.6	1821760.6	8504482.0

**Tabela 6.1.1.7: Združeni LFS-podatki: Belgija 2002 – regija**

Regija	N
1 Brussels	800816.4
2 Flanders	4966557.1
3 Wallonia	2737108.4
Skupaj	8504482.0

V meritvi 2 ima Belgija (ESS DATA 2005c: 4–13) predstavljenih kar nekaj podatkov: spol/starost, izobrazba, spol/starost/izobrazba (samo za starost 15–64 let) in spol/regija. Vsak sklop podatkov je predstavljen najprej kot populacijsko število prebivalcev in njihovi odstotki, enako tudi za vzorec. Sledi primerjava odstotkov med populacijo in vzorcem, kar je tudi prikazano na stolpcnem grafu. Tu bom predstavila samo populacijsko število prebivalcev za sklope, ki sem jih uporabila pri sestavljanju tabele, ustrezne za uteževanje. Osnovna vira predstavljenih podatkov sta dva, Eurostat (spol/starost) in belgijski statistični urad (ostale).

**Tabela 6.1.1.8: Populacija: Belgija R2 spol/starost**

Age	male	female	total
15–19	311969	299048	611017
20–24	325731	319275	645006
25–29	328486	323553	652039
30–34	370439	361880	732319
35–39	400179	388537	788716
40–44	408212	400296	808508
45–49	382008	377481	759489
50–54	349780	345395	695175
55–59	318127	320295	638422

Age	male	female	total
60–64	237637	250534	488171
65–70	232610	261357	493967
70–74	210246	262916	473162
75–79	156109	228245	384354
80–84	95329	171851	267180
85–89	28450	70148	98598
90–94	11411	39416	50827
Total	4166723	4420227	8586950

Vir: ESS DATA 2005c: 4.

**Tabela 6.1.1.9: Populacija: Belgija R2 izobrazba**

	N
Lower	2074754
Lower secondary	1843233
Higher secondary	2695753
Post-secondary	1328881
University and postgraduate	656026
Total	8598647

Vir: ESS DATA 2005c: 7.

**Tabela 6.1.1.10: Populacija: Belgija R2 regija samo NUTS1**

	Male	Female	Total
Flemish region	2966640	3049384	6016024
Brussels region	480334	519565	999899
Walloon region	1640202	1740296	3380498
Total	5087176	5309245	10396421

Vir: ESS DATA 2005c: 12.

Tabela za regijo je brez NUTS 2 vrednosti že ustrezna za uteževanje. Ostali dve tabeli preoblikujem kot pri R1:

**Tabela 6.1.1.11: Združeni podatki iz ESS, Belgija R2, primerni za raking (spol/starost)**

	15–34	35–54	55+	Skupaj
moški	1336625	1540179	1289919	4166723
ženske	1303756	1511709	1604762	4420227
skupaj	2640381	3051888	2894681	8586950

**Tabela 6.1.1.12: Združeni podatki iz ESS, Belgija R2, primerni za raking (izobrazba)**

Izobrazba	N
nizka (1 + 2)	3917987
srednja (3)	2695753
visoka (4 + 5)	1984907
skupaj	8598647

V vseh treh tabelah se skupna vsota (*angl. total*) razlikuje. V tabeli za regijo je najvišja, verjetno so vključili vse prebivalstvo. Razlika med drugima dvema tabelama je morda ta, da sta vzeti iz različnih virov. Lahko bi sicer dopolnili tabelo spol/starost z manjkajočimi vrednostmi, a zaradi manjše primerljivosti to ni najboljše. Podatki iz LFS so obravnavani na enak način kot v meritvi 1 in so primerni za uteževanje.

**Tabela 6.1.1.13: Združeni LFS-podatki: Belgija 2004 – spol/starost/izobrazba**

spol	starost	izobrazba			skupaj
		nizka	srednja	visoka	
moški	15–34	491284.0	544182.8	300696.1	1336162.8
	35–54	574204.2	529948.7	435987.0	1540140.0
	55+	768111.7	310689.9	212887.3	1291689.0
	skupaj	1833599.9	1384821.4	949570.4	4167991.8
ženske	15–34	400856.7	520149.8	382201.9	1303208.4
	35–54	544915.2	519756.8	446994.9	1511667.0
	55+	1175084.0	281863.5	157897.6	1614845.2
	skupaj	2120855.9	1321770.2	987094.4	4429720.6
skupaj		3954455.9	2706591.6	1936664.9	8597712.3

**Tabela 6.1.1.14: Združeni LFS-podatki: Belgija 2004 – regija**

Regija	N
1 Brussels	817541.3
2 Flanders	5017435.1
3 Wallonia	2762735.9
Skupaj	8597712.3

V meritvi 3 belgijski nacionalni koordinatorji priskrbijo tri populacijske tabele:

**Tabela 6.1.1.15: Populacija: Belgija R3 spol/starost**

Age	male	female
15–24	639177	621771
25–34	693742	681517
35–44	804116	784595
45–54	741535	734419
55–64	569017	581605
65–74	443561	522181
75 plus	302686	531072
total	4193834	4457160

Vir: ESS DATA 2007c: 4

**Tabela 6.1.1.16: Populacija: Belgija R3 izobrazba**

	N
Lower	1989662
Lower secondary	1827658
Higher secondary	2811260
Post-secondary	1381291
University and postgraduate	670618
Total	8680489

Vir: ESS DATA 2007c: 5

Tabela za regijo je brez vrednosti NUTS 2 že ustrezena za uteževanje, popravila sem le tretjo regijo, *Walloon region*, saj je v originalu število prebivalcev za to regijo 10541893, kar je v resnici skupna vsota.

**Tabela 6.1.1.17: Populacija: Belgija R3 regija, samo NUTS1 popravljena**

	Male	Female	Total
Flemish region	3006244	3089172	6095416
Brussels region	492994	531498	1024492
Walloon region	1660709	1761276	3421985
Total	5159947	5381946	10541893

Vir: ESS DATA 2007c: 4

Seštevek NUTS 2 regij nam da vsoto, ki je v zgornji tabeli. Poglejmo si še preoblikovani tabeli spol/starost in izobrazba.

**Tabela 6.1.1.18: Združeni podatki iz ESS-priloge, primerni za raking (spol/starost)**

	15–34	35–54	55+	Skupaj
moški	1332919	1545651	1315264	4193834
ženske	1303288	1519014	1634858	4457160
Skupaj	2636207	3064665	2950122	8650994

**Tabela 6.1.1.19: Združeni podatki iz ESS-priloge, primerni za raking (izobrazba)**

Izobrazba	N
Nizka (1 + 2)	3817320
Srednja (3)	2811260
Visoka (4 + 5)	2051909
Skupaj	8680489

LFS-podatki so podobni kot v prejšnjih dveh meritvah in tako ustrezni za uteževanje.

**Tabela 6.1.1.20: Združeni LFS-podatki: Belgija 2006 – spol/starost/izobrazba**

spol	starost	izobrazba				skupaj
		nizka	srednja	visoka	manjkajoča	
moški	15–34	466678.4	563880.1	304101.1	0.0	1334659.7
	35–54	524903.2	564633.7	469634.3	0.0	1559171.2
	55+	768233.7	326070.2	255706.6	0.0	1350010.5
	skupaj	1759815.3	1454584.0	1029442.0	0.0	4243841.3
ženske	15–34	389190.6	515493.5	403963.4	207.0	1308854.5
	35–54	491996.4	554834.8	485418.7	0.0	1532249.9
	55+	1123096.3	325687.3	215631.2	0.0	1664414.8
	skupaj	2004283.4	1396015.5	1105013.3	207.0	4505519.3
skupaj		3764098.7	2850599.5	2134455.4	207.0	8749360.6

**Tabela 6.1.1.21: Združeni LFS- podatki: Belgija 2006 – regija**

Regija	N
1 Brussels	835237.0
2 Flanders	5102866.0
3 Wallonia	2811257.6
Skupaj	8749360.6

V naslednji tabeli so prikazani odstotki za kategorije. Kot lahko vidimo, ni večjih odstopanj med ESS- in LFS- populacijskimi podatki. Ker pa so v ESS ti podatki predstavljeni bolj razdrobljeno, je bolje uporabiti LFS-podatke za uteževanje.

**Tabela 6.1.1.22: Primerjava ESS in LFS (%), skupine za spol, starost in izobrazbo, Belgija**

BE meritev	vir	spol		starost			samo 65+	izobrazba			
		m	ž	15–34	35–54	55+		1	2	3	neznano
R1 2002	ESS	48.44	51.56	31.20	35.61	33.19	20.53	48.56	29.86	21.15	0.43
	LFS	48.44	51.56	31.20	35.61	33.19		48.17	30.41	21.42	/
R2 2004	ESS	48.52	51.48	30.75	35.54	33.71	20.59	45.57	31.35	23.08	
	LFS	48.48	51.52	30.70	35.50	33.81		45.99	31.48	22.53	/
R3 2006	ESS	48.48	51.52	30.47	35.43	34.10	20.80	43.98	32.39	23.64	/
	LFS	48.50	51.50	30.21	35.33	34.45		43.02	32.58	24.40	0.00

## 6.1.2 Slovenija

Populacijske podatke za Slovenijo je naš nacionalni koordinator vzel iz cenzusa 2002. Za meritev 1 so v ESS (ESS DATA 2003d: 70–75) na voljo tabela spol/starost/regija za prebivalstvo vseh starosti in tabeli izobrazba/starost/spol ter izobrazba/spol/regija za prebivalce nad 15 let, ki sta predstavljeni tudi spodaj.

**Tabela 6.1.2.1: Populacijski podatki: Slovenija – spol/starost/izobrazba**

	sum	Education								
		0 no education	1 incomplete primary	2 primary	secondary			5 high school	university	
					sum	3 vocational school	4 general secondary		6 Under-graduate	7 post graduate
<b>Sum</b>										
<b>Sum</b>	<b>1663869</b>	<b>11337</b>	<b>104219</b>	<b>433910</b>	<b>899341</b>	<b>452292</b>	<b>447049</b>	<b>84044</b>	<b>114630</b>	<b>16388</b>
15–19	130029	530	1328	104645	23508	11628	11880	13	5	-
20–24	147687	554	1207	14400	127666	39863	87803	964	2880	16
25–29	144977	478	1465	15450	98565	43463	55102	6676	21233	1110
30–34	140612	518	1472	22005	87353	43052	44301	8799	18166	2299
35–39	153518	549	1886	23536	99462	50667	48795	11107	14490	2488
40–44	152142	673	4224	28394	91558	49537	42021	11573	13422	2298
45–49	158611	892	7639	36189	89553	51403	38150	10394	12001	1943
50–54	138542	875	8450	32718	74776	41168	33608	9936	9886	1901
55–59	105161	737	8673	24615	55675	31788	23887	7740	6411	1310
60–64	103609	908	12977	29053	48152	28990	19162	5835	5616	1068
65 +	288981	4623	54898	102905	103073	60733	42340	11007	10520	1955
<b>Males</b>										
<b>Sum</b>	<b>804286</b>	<b>4092</b>	<b>42400</b>	<b>169509</b>	<b>487288</b>	<b>280373</b>	<b>206915</b>	<b>36083</b>	<b>55070</b>	<b>9844</b>
15–19	66417	z	897	53168	12032	6996	5036	8	z	-
20–24	76190	z	851	9198	64415	24901	39514	456	z	8
25–29	74456	259	1057	8963	53713	26984	26729	2494	7429	541
30–34	71306	285	1012	10913	47829	26836	20993	3104	7020	1143
35–39	77366	278	1144	10695	53777	31667	22110	3784	6326	1362
40–44	77125	294	2281	11852	50523	31269	19254	4360	6508	1307
45–49	82023	395	4237	14435	51299	33012	18287	4310	6226	1121
50–54	71061	369	4569	11775	43003	26679	16324	4528	5635	1182
55–59	51517	268	3983	7723	31082	19986	11096	3615	3921	925
60–64	49407	307	5528	8275	27780	18585	9195	3086	3667	764
65 +	107418	1001	16841	22512	51835	33458	18377	6338	7400	1491
<b>Females</b>										
<b>Sum</b>	<b>859583</b>	<b>7245</b>	<b>61819</b>	<b>264401</b>	<b>412053</b>	<b>171919</b>	<b>240134</b>	<b>47961</b>	<b>59560</b>	<b>6544</b>
15–19	63612	z	431	51477	11476	4632	6844	5	z	-
20–24	71497	z	356	5202	63251	14962	48289	508	z	8
25–29	70521	219	408	6487	44852	16479	28373	4182	13804	569
30–34	69306	233	460	11092	39524	16216	23308	5695	11146	1156
35–39	76152	271	742	12841	45685	19000	26685	7323	8164	1126
40–44	75017	379	1943	16542	41035	18268	22767	7213	6914	991
45–49	76588	497	3402	21754	38254	18391	19863	6084	5775	822
50–54	67481	506	3881	20943	31773	14489	17284	5408	4251	719
55–59	53644	469	4690	16892	24593	11802	12791	4125	2490	385
60–64	54202	601	7449	20778	20372	10405	9967	2749	1949	304
65 +	181563	3622	38057	80393	51238	27275	23963	4669	3120	464

Vir: ESS DATA 2003d: 72–73.

Pri pripravi tabele, ustrezne za uteževanje, je potrebno opozoriti na vmesni stolpec sum, ki predstavlja celotno število sekundarno izobraženih in se ne upošteva pri združevanju kategorij.

**Tabela 6.1.2.2: Populacijski podatki: Slovenija – izobrazba/spol/regija**

Regional Unit	Sum	Education								
		no education	incomplete primary	primary	secondary			general secondary	university	
					sum	vocational school	high school		under-graduate	post graduate
<b>SLOVENIJA</b>	<b>1663869</b>	<b>11337</b>	<b>104219</b>	<b>433910</b>	<b>899341</b>	<b>452292</b>	<b>447049</b>	<b>84044</b>	<b>114630</b>	<b>16388</b>
Pomurska	103272	850	5272	39189	49745	28508	21237	3920	3923	373
Podravska	266296	1727	14640	69628	149728	77537	72191	13567	14923	2083
Koroška	61597	464	5119	15887	34108	18952	15156	2787	2955	277
Savinjska	213604	1703	13294	60567	116410	62351	54059	9729	10853	1048
Zasavska	38932	207	2695	11058	21306	11743	9563	1719	1805	142
Spodnjeposavska	57892	480	5301	16179	30678	16563	14115	2312	2693	249
Jugovzhodna	112916	1183	11608	31458	57478	30929	26549	4779	5820	590
Osrednjeslovenska	412782	2089	17503	88214	227563	99405	128158	24452	44213	8748
Gorenjska	164050	966	10242	40907	90379	44905	45474	8462	11658	1436
Notranjsko-kraška	42737	351	3765	11914	22139	11362	10777	2094	2281	193
Goriška	101112	609	7498	29083	51886	27165	24721	4849	6606	581
Obalno-kraška	88679	708	7282	19826	47921	22872	25049	5374	6900	668
<b>males</b>	<b>804286</b>	<b>4092</b>	<b>42400</b>	<b>169509</b>	<b>487288</b>	<b>280373</b>	<b>206915</b>	<b>36083</b>	<b>55070</b>	<b>9844</b>
Pomurska	49851	310	2535	15171	27901	17894	10007	1769	1937	228
Podravska	128483	678	6126	25924	81240	47550	33690	5967	7198	1350
Koroška	30491	190	2263	6368	19011	12039	6972	1178	1321	160
Savinjska	103696	590	5798	23877	63595	38630	24965	4079	5122	635
Zasavska	18711	61	1015	4063	11918	7352	4566	734	837	83
Spodnjeposavska	28192	180	2223	6362	17037	10151	6886	976	1271	143
Jugovzhodna	55239	454	5154	12633	31861	19446	12415	2084	2717	336
Osrednjeslovenska	196692	769	6458	35437	117700	60294	57406	10065	21143	5120
Gorenjska	79202	310	3705	15894	49008	27520	21488	3607	5738	940
Notranjsko-kraška	20979	108	1523	4741	12472	7381	5091	903	1119	113
Goriška	49447	239	3057	10990	29534	17926	11608	2078	3201	348
Obalno-kraška	43303	203	2543	8049	26011	14190	11821	2643	3466	388
<b>females</b>	<b>859583</b>	<b>7245</b>	<b>61819</b>	<b>264401</b>	<b>412053</b>	<b>171919</b>	<b>240134</b>	<b>47961</b>	<b>59560</b>	<b>6544</b>
Pomurska	53421	540	2737	24018	21844	10614	11230	2151	1986	145
Podravska	137813	1049	8514	43704	68488	29987	38501	7600	7725	733
Koroška	31106	274	2856	9519	15097	6913	8184	1609	1634	117
Savinjska	109908	1113	7496	36690	52815	23721	29094	5650	5731	413
Zasavska	20221	146	1680	6995	9388	4391	4997	985	968	59
Spodnjeposavska	29700	300	3078	9817	13641	6412	7229	1336	1422	106
Jugovzhodna	57677	729	6454	18825	25617	11483	14134	2695	3103	254
Osrednjeslovenska	216090	1320	11045	52777	109863	39111	70752	14387	23070	3628
Gorenjska	84848	656	6537	25013	41371	17385	23986	4855	5920	496
Notranjsko-kraška	21758	243	2242	7173	9667	3981	5686	1191	1162	80
Goriška	51665	370	4441	18093	22352	9239	13113	2771	3405	233
Obalno-kraška	45376	505	4739	11777	21910	8682	13228	2731	3434	280

Vir: ESS DATA 2003d: 74–75.

V meritvi 2 (ESS DATA 2005c: 120–126) so predstavljene identične tabele, na podlagi cenzusa 2002, ni pa zgornje tabele za izobrazbo/starost/regijo. Slednja je ponovno predstavljena za R3 (ESS DATA 2007c: 140–141), kjer pa ni ostalih tabel. Tako imamo za vse tri meritve enake podatke in lahko prikažem samo dve tabeli (spol/starost/izobrazba in regija), ki sta ustrezni za uteževanje. Identičnost podatkov v vseh treh meritvah pa ni najbolj primerna, saj se prebivalstvo med leti spreminja.

Podatke iz tabel lahko združimo v tabelo spol/starost/izobrazba, a se zaradi neprikazanih enot (celice, ki vsebujejo črko z) skupna vsota ne ujema, zato ni primerna za uteževanje, saj izpuščamo določene enote.

**Tabela 6.1.2.3: Združeni ESS-podatki: Slovenija 2002 – spol/starost/izobrazba**

spol	starost	izobrazba			skupaj
		nizka (0 + 1 + 2)	srednja (3 + 4 + 5)	visoka (6+7)	
moški	15–34	86603	184051	16141	286795
	35–54	62324	215584	29667	307575
	55+	66438	123736	18168	208342
	skupaj	215365	523371	63976	802712
ženske	15–34	76365	169493	26683	272541
	35–54	83701	182775	28762	295238
	55+	172951	107746	8712	289409
	skupaj	333017	460014	64157	857188
skupaj		548382	983385	128133	1659900

LFS-jeve populacijske podatke imamo za vsa tri leta in so primerni za uteževanje, le predhodno je potrebno na vzorcu združiti kategorije za regijo, ki je NUTS 3, po kategorijah, opisanih v Tabela 5.3.4.2, v nivo NUTS 2, ki je na populaciji.

**Tabela 6.1.2.4: Združeni LFS-podatki: Slovenija 2002 – spol/starost/izobrazba**

spol	starost	izobrazba				skupaj
		nizka	srednja	visoka	manjkajoča	
moški	15–34	87941.3	187849.3	19683.8	2779.4	298253.7
	35–54	56874.1	206536.4	38278.6	5374.4	307063.6
	55+	64663.8	117299.7	26710.2	4121.2	212794.9
	skupaj	209479.2	511685.4	84672.7	12275.0	818112.2
ženske	15–34	75400.9	160271.2	36155.1	6998.4	278825.4
	35–54	83499.3	166138.3	43696.2	6139.1	299472.9
	55+	165993.6	104065.8	18619.1	2203.7	290882.3
	skupaj	324893.8	430475.3	98470.4	15341.2	869180.6
skupaj		534372.9	942160.7	183143.0	27616.1	1687292.8

**Tabela 6.1.2.5: Združeni LFS-podatki: Slovenija 2002 – regija**

Regija	N
1 Vzhodna Slovenija	913370.8
2 Zahodna Slovenija	773922.0
Skupaj	1687292.8

**Tabela 6.1.2.6: Združeni LFS-podatki: Slovenija 2004 – spol/starost/izobrazba**

spol	starost	izobrazba				skupaj
		nizka	srednja	visoka	manjkajoča	
moški	15–34	82666.1	185041.4	26695.6	552.1	294955.2
	35–54	59138.0	203698.1	48568.6	0.0	311404.7
	55+	58560.1	126354.3	36729.8	0.0	221644.2
	skupaj	200364.2	515093.8	111994.0	552.1	828004.1
ženske	15–34	70553.2	158550.1	49666.8	348.7	279118.8
	35–54	74250.1	169006.0	58065.8	0.0	301321.9
	55+	163957.6	107171.6	26808.7	0.0	297937.9
	skupaj	308760.9	434727.8	134541.2	348.7	878378.6
skupaj		509125.1	949821.6	246535.2	900.8	1706382.7

**Tabela 6.1.2.7: Združeni LFS-podatki: Slovenija 2004 – regija**

Regija	N
1 Vzhodna Slovenija	922995.1
2 Zahodna Slovenija	783387.6
Skupaj	1706382.7

**Tabela 6.1.2.8: Združeni LFS-podatki: Slovenija 2006 – spol/starost/izobrazba**

spol	starost	izobrazba			skupaj
		nizka	srednja	visoka	
moški	15–34	73122.0	183770.1	30207.7	287099.7
	35–54	53895.5	203510.5	58011.9	315417.9
	55+	64561.4	130307.6	41083.9	235952.9
	skupaj	191578.9	517588.1	129303.5	838470.5
ženske	15–34	62013.6	151802.3	53927.7	267743.6
	35–54	65103.4	169209.0	70972.3	305284.7
	55+	158333.4	121694.0	32081.6	312109.0
	skupaj	285450.4	442705.3	156981.6	885137.3
skupaj		477029.2	960293.5	286285.1	1723607.9

**Tabela 6.1.2.9: Združeni LFS-podatki: Slovenija 2006 – regija**

Regija	N
1 Vzhodna Slovenija	929046.3
2 Zahodna Slovenija	794561.6
Skupaj	1723607.9

V naslednji tabeli so prikazani odstotki za kategorije. Kot lahko vidimo, ni večjih odstopanj med ESS- in LFS- populacijskimi podatki za spol in starost. Večja razlika je edino pri izobrazbi, a se izobrazbenih razredov glede na ESS-jeve

populacijske kategorije ne da združiti na bolj primerljiv način. Bolje je uporabiti LFS-podatke, saj zajamejo vse enote, poleg tega se skozi leta spreminja, kot se spreminja tudi prebivalstvo Slovenije (opazimo lahko npr. trend naraščanja starega prebivalstva), ESS pa za vse tri meritve vsebuje enake podatke.

**Tabela 6.1.2.10: Primerjava ESS in LFS (%), skupine za spol, starost in izobrazbo, Slovenija**

Sl meritev	vir	spol		starost			samo 65+	izobrazba			
		m	ž	15–34	35–54	55+		1	2	3	neznano
R1 2002	ESS	48.34	51.66	33.86	36.23	29.92	17.37	33.02	59.10	7.87	/
	LFS	48.49	51.51	34.20	35.95	29.85		31.67	55.84	10.85	1.64
R2 2004	ESS	48.34	51.66	33.86	36.23	29.92	17.37	33.02	59.10	7.87	/
	LFS	48.52	51.48	33.64	35.91	30.45		29.84	55.66	14.45	0.05
R3 2006	ESS	48.34	51.66	33.86	36.23	29.92	17.37	/	/	/	/
	LFS	48.65	51.35	32.19	36.01	31.80		27.68	55.71	16.61	/

### 6.1.3 Švedska

Populacijskih podatkov za meritev 1 in meritev 2 ni na voljo v ESS-prilogi. V tretji meritvi so med populacijskimi podatki (ESS DATA 2007c: 134–135) samo vzorčni podatki za spol, starost, regijo in tri druge spremenljivke, a nam ne koristijo, ker to niso pravi populacijski podatki.

LFS-jeve populacijske podatke imamo za vsa tri leta in so primerni za uteževanje, tudi regija se ujema z regijo na vzorcu (enake kategorije). Pojavijo se edino manjkajoče vrednosti za regijo, ki jih na vzorcu ni. Le-te obravnavam v naslednjem poglavju.

**Tabela 6.1.3.1: Združeni LFS-podatki: Švedska 2002 – spol/starost/izobrazba**

spol	starost	izobrazba					skupaj
		nizka	srednja	visoka	manjkajoča		
moški	15–34	251882.5	604110.1	183530.7	97241.7	1136765.0	
	35–54	243715.8	727660.9	263739.2	8196.9	1243312.8	
	55+	356335.0	370564.0	165907.0	311625.0	1204430.8	
	skupaj	851933.3	1702334.9	613176.8	417063.5	3584508.6	
ženske	15–34	227225.1	543561.1	227618.6	94335.5	1092740.3	
	35–54	167950.9	660092.8	364114.2	8351.7	1200509.6	
	55+	354579.0	376696.9	196026.5	495773.3	1423075.8	
	skupaj	749755.0	1580350.8	787759.3	598460.5	3716325.6	
	skupaj	1601688.3	3282685.7	1400936.2	1015524.1	7300834.2	

**Tabela 6.1.3.2: Združeni LFS-podatki: Švedska 2002 – regija**

Regija	N
0	792315.0
1 Stockholm	1317459.5
2 Östra Mellansverige	1092035.5
4 Sydsverige	957452.3
6 Norra Mellansverige	616953.4
7 Mellersta Norrland	288790.4
8 Övre Norrland	390953.5
9 Småland med öarna	580038.7
0A Västsverige	1264835.9
Skupaj	7300834.2

**Tabela 6.1.3.3: Združeni LFS-podatki: Švedska 2004 – spol/starost/izobrazba**

spol	starost	izobrazba				skupaj
		nizka	srednja	visoka	manjkajoča	
moški	15–34	258839.4	578558.9	195483.0	110587.4	1143468.7
	35–54	217856.6	740963.5	273738.3	8730.7	1241289.0
	55+	362116.7	400104.1	174162.7	311347.5	1247731.0
	skupaj	838812.6	1719626.4	643384.0	430665.7	3632488.7
ženske	15–34	235670.8	506255.9	256899.2	99185.3	1098011.2
	35–54	149902.3	656375.6	387230.8	5460.5	1198969.2
	55+	335101.2	411598.7	216612.8	494861.2	1458173.8
	skupaj	720674.2	1574230.2	860742.8	599507.0	3755154.2
skupaj		1559486.8	3293856.6	1504126.8	1030172.7	7387642.9

**Tabela 6.1.3.4: Združeni LFS-podatki: Švedska 2004 – regija**

Regija	N
0	791511.2
1 Stockholm	1325359.4
2 Östra Mellansverige	1116549.1
4 Sydsverige	958158.8
6 Norra Mellansverige	629847.3
7 Mellersta Norrland	297928.0
8 Övre Norrland	399110.6
9 Småland med öarna	592393.9
0A Västsverige	1276784.6
Skupaj	7387642.9

**Tabela 6.1.3.5: Združeni LFS-podatki: Švedska 2006 – spol/starost/izobrazba**

spol	starost	izobrazba				skupaj
		nizka	srednja	visoka	manjkajoča	
moški	15–34	289707.6	545170.7	217945.7	113690.0	1166514.0
	35–54	196458.8	745446.6	295648.9	5301.4	1242855.6
	55+	326210.8	436414.8	208538.9	322982.1	1294146.5
	skupaj	812377.1	1727032.1	722133.5	441973.4	3703516.1
ženske	15–34	254567.8	481094.5	282463.6	105303.4	1123429.2
	35–54	134117.4	650343.2	413271.8	5098.2	1202830.6
	55+	312221.9	439371.8	246231.7	495798.1	1493623.5
	skupaj	700907.1	1570809.5	941967.1	606199.6	3819883.3
skupaj		1513284.2	3297841.7	1664100.6	1048173.0	7523399.4

**Tabela 6.1.3.6: Združeni LFS-podatki: Švedska 2006 – regija**

Regija	N
0	799100.8
1 Stockholm	1422481.8
2 Östra Mellansverige	1125405.4
4 Sydsverige	982495.0
6 Norra Mellansverige	607814.1
7 Mellersta Norrland	272734.0
8 Övre Norrland	381948.7
9 Småland med öarna	586398.6
0A Västsverige	1345021.1
Skupaj	7523399.4

Tu za uteževanje uporabimo populacijske podatke LFS, saj drugih ni na voljo.

**Tabela 6.1.3.7: Primerjava ESS in LFS (%), skupine za spol, starost in izobrazbo, Švedska**

SE meritev	vir	spol		starost			samo 65+	izobrazba			
		m	ž	15–34	35–54	55+		1	2	3	neznano
R1 2002	ESS	/	/	/	/	/	/	/	/	/	/
	LFS	49.10	50.90	30.54	33.47	35.99	20.97	21.94	44.96	19.19	13.91
R2 2004	ESS	/	/	/	/	/	/	/	/	/	/
	LFS	49.17	50.83	30.34	33.03	36.63	20.76	21.11	44.59	20.36	13.94
R3 2006	ESS	/	/	/	/	/	/	/	/	/	/
	LFS	49.23	50.77	30.44	32.51	37.05	20.90	20.11	43.83	22.12	13.93

## 6.2 Problem manjkajočih vrednosti v podatkih za uteževanje

Tu bom obravnavala manjkajoče podatke, ki so posledica neodgovora na kontrolne spremenljivke. V splošnem se neodgovor spremenljivke najpogosteje obravnava z enostavnimi postopki vstavljanja, mogoče ga je reševati z metodami neposredne analize, v praksi gre največkrat za enostavne postopke, ki vstavijo vrednosti za manjkajoče odgovore s pomočjo odgovorov na ostala vprašanja. Ponekod lahko vstavimo aritmetično sredino, ki smo jo izračunali pri respondentih z zanimi vrednostmi, a tako umečno spremenimo porazdelitev spremenljivke ipd. (Kalton in Vehovar 2001: 89). Tu prikazujem še malce drugačen način prirejanja manjkajočih vrednosti, saj lahko podatki manjkajo na populaciji ali na vzorcu, bolj specifično za vsako spremenljivko, ki je uporabljen pri uteževanju (raking na tabeli spol/starost/izobrazba in regija).

Osredotočila sem se na tri države (Belgijo, Slovenijo in Švedsko). Vzorčne podatke sem pridobila s strani ESS DATA (2003a, 2005a in 2007a). Najprej pogledam populacijski tabeli (spol/starost/izobrazba in regija): Vsebujeta kakšne manjkajoče vrednosti? Kakšen je odstotek le-teh? Nato pogledam tabeli z enakimi spremenljivkami na vzorcu, če vsebujejo manjkajoče vrednosti, s tem da sem prej rekodirala spremenljivke, tako da manjkajoče vrednosti predstavljajo vrednost 0. Glede na postavitev manjkajočih vrednosti sem določila 5 glavnih situacij, ki so opisane spodaj. Vsaki situaciji sledi določen postopek prilagajanja podatkov na manjkajoče vrednosti (reševanje problema).

Postopki reševanja problema (P) za dane situacije (S) so naslednji:

S1: manjkajo podatki za eno populacijsko spremenljivko (npr. izobrazba). Le-teh je manj kot 1 %. Na vzorcu so vsi podatki za kontrolne spremenljivke.

P1: glede na izobrazbeno porazdelitev ostalih skupin te manjkajoče vrednosti porazdelim med ostale izobrazbene vrednosti.

S2: manjkajo podatki za eno populacijsko spremenljivko (npr. izobrazba). Le-teh je več kot 1 %. Na vzorcu so vsi podatki za kontrolne spremenljivke.

P2: prilagajam na dva načina – prvi enak kot P1, v drugem pa manjkajoče vrednosti preprosto izbrišem. Za drugo tabelo (regijo) iz primarnih populacijskih podatkov sestavim tabelo z vsemi štirimi spremenljivkami, pobrišem manjkajoče vrednosti za izobrazbo in ponovno pridobim vrednosti za regijo. Za končno uteževanje se kasneje uporabi način, kjer je pristranskost manjša.

S3: manjkajo podatki za dve populacijski spremenljivki (izobrazba in regija). Na vzorcu so vsi podatki za kontrolne spremenljivke.

P3: lahko ponovim P1 in P2 postopka, vsakega posebej za eno spremenljivko (odvisno od višine % za manjkajoče vrednosti). Ker sta regija in izobrazba v različnih tabelah, druga na drugo ne vplivata.

S4: na vzorcu so manjkajoče vrednosti za eno ali več spremenljivk. Populacijska tabela nima manjkajočih vrednosti.

P4: izračunam odstotke za te manjkajoče vrednosti na vzorcu, tako da se odstotki vrednosti za spremenljivke, kjer ni manjkajočih vrednosti, ne spremenijo. Odstotke na enak način izračunam za populacijo. Vstavim odstotke manjkajočih vrednosti v populacijo in preračunam prejšnje populacijske odstotke.

S5: mankajo tako podatki na populaciji kot tudi na vzorcu in se ne nahajajo v istih celicah.

P5: najprej sledim postopkom o manjkajočih vrednostih na populaciji. Ko dobim populacijsko tabelo brez manjkajočih vrednosti, se lotim še postopka na vzorcu.

Seveda obstaja še več pod situacij glede na spremenljivke, ki imajo manjkajoče vrednosti. Kadar so manjkajoče vrednosti na vzorcu, izračunamo odstotke za celice, kjer se nahajajo manjkajoče vrednosti, način računanja odstotkov pa je odvisen od situacije. V naslednji tabeli so predstavljene situacije za tri izbrane države v ESS-meritvah.

**Tabela 6.1.3.1: Pregled situacij v ESS za Belgijo, Slovenijo in Švedsko**

	R1	R2	R3
Belgia	S4, spol, starost in izobrazba	S4, starost in izobrazba	S5 (S1, izobrazba)
Slovenija	S5 (S2, S3, spol, izobrazba)	S5 (S1, spol, starost, izobrazba)	S4, starost in izobrazba
Švedska	S5 (S3, izobrazba, regija, S4, spol, starost in izobrazba)	S5 (S3, izobrazba in regija, S4, starost in izobrazba)	S5 (S3 izobrazba, regija, S4, spol, starost in izobrazba)

Ker se situacije ponavljajo, bom predstavila pet primerov, ki so v Tabela 6.1.3.1 pobarvani z viola. Pred začetkom pa naj še enkrat omenim, da imajo ESS-jeve vzorčne baze še dve uteži, ki jih je treba upoštevati vnaprej. Ena je vzorčna utež (*dweight*), ki je za izbrane tri države povsod enaka 1, torej nima vpliva na število enot<sup>4</sup>; druga je populacijska utež (*pweight*), ki se jo uporabi takrat, kadar delamo primerjavo med državami. Ker tu predstavljam rezultate za vsako državo posebej, utež *pweight* ni uporabljena.

---

<sup>4</sup> Če *dweight* ni enaka 1, je potrebno vzorčno bazo najprej utežiti s to utežjo in za vzorčno tabelo uporabiti to, uteženo.

## 6.2.1 Belgija R1

Najprej pogledamo tabeli za populacijo (Tabela 6.1.1.6 in Tabela 6.1.1.7) ter vzorec (Tabela 6.2.1.1 in Tabela 6.2.1.2):

**Tabela 6.2.1.1: Vzorec spol/starost/izobrazba (n), Belgija R1**

spol	starost	izobrazba				skupaj
		0 manjkajoča	1 nizka	2 srednja	3 visoka	
neznano	0	8	5	2	2	17
	15–34	0	0	3	2	5
	35–54	0	1	0	1	2
	55+	0	4	1	0	5
	skupaj	8	10	6	5	29
moški	0	3	5	13	15	36
	15–34	3	78	144	94	319
	35–54	5	91	147	94	337
	55+	1	133	79	58	271
	skupaj	12	307	383	261	963
ženske	0	1	9	10	14	34
	15–34	4	69	118	80	271
	35–54	2	102	121	118	343
	55+	3	158	61	37	259
	skupaj	10	338	310	249	907
skupaj		30	655	699	515	1899

**Tabela 6.2.1.2: Vzorec regija (n), Belgija R1**

Regija	n
1 Flemish region	1234
2 Brussels region	90
3 Walloon region	575
Skupaj	1899

Na populaciji ni manjkajočih vrednosti, na vzorcu imamo manjkajoče vrednosti za spol, starost in izobrazbo. Regijska populacijska tabela ostane enaka, tabela spol/starost/izobrazba pa potrebuje prilagajanje. Ker imajo vse tri spremenljivke manjkajoče vrednosti, odstotke izračunamo glede na celoto. To naredimo tako za populacijo kot tudi za vzorec:

**Tabela 6.2.1.3: Populacija spol/starost/izobrazba (%), Belgija 2002**

spol	starost	izobrazba			skupaj
		nizka	srednja	visoka	
moški	15–34	6.02	6.53	3.27	15.82
	35–54	7.15	5.88	4.95	17.98
	55+	9.31	3.02	2.32	14.65
	skupaj	22.48	15.43	10.53	48.44
ženske	15–34	4.95	6.15	4.29	15.38
	35–54	7.11	5.70	4.82	17.63
	55+	13.63	3.13	1.78	18.54
	skupaj	25.69	14.98	10.89	51.56
skupaj		48.17	30.41	21.42	100.00

**Tabela 6.2.1.4: Vzorec spol/starost/izobrazba (%), Belgija R1**

spol	starost	izobrazba				skupaj
		0 manjkajoča	1 nizka	2 srednja	3 visoka	
neznano	0	0.42	0.26	0.11	0.11	0.90
	15–34	0.00	0.00	0.16	0.11	0.26
	35–54	0.00	0.05	0.00	0.05	0.11
	55+	0.00	0.21	0.05	0.00	0.26
	skupaj	0.42	0.53	0.32	0.26	1.53
moški	0	0.16	0.26	0.68	0.79	1.90
	15–34	0.16	4.11	7.58	4.95	16.80
	35–54	0.26	4.79	7.74	4.95	17.75
	55+	0.05	7.00	4.16	3.05	14.27
	skupaj	0.63	16.17	20.17	13.74	50.71
ženske	0	0.05	0.47	0.53	0.74	1.79
	15–34	0.21	3.63	6.21	4.21	14.27
	35–54	0.11	5.37	6.37	6.21	18.06
	55+	0.16	8.32	3.21	1.95	13.64
	skupaj	0.53	17.80	16.32	13.11	47.76
skupaj		1.58	34.49	36.81	27.12	100.00

Celice, kjer se pojavijo manjkajoče vrednosti na vzorcu (obarvane s sivo), dodamo populacijski tabeli, a pri tem prilagodimo obstoječe populacijske celice, tako da razmerja ostanejo enaka: odstotek v populaciji množimo z vsoto vseh nemanjkajočih na vzorcu (100 – 6.16). 6.16 % je vsota vseh manjkajočih vrednosti na vzorcu. **Primer izračuna:** celica moški, starost 15–34 let, nizka izobrazba:  $6.02 * (100 \% - 6.16 \%) = 5.65$ .

**Tabela 6.2.1.5: Dopolnjena populacija spol/starost/izobrazba (%), Belgija 2002/R1**

spol	starost	izobrazba				skupaj
		1 nizka	2 srednja	3 visoka	0 manjkajoča	
moški	15–34	5.65	6.13	3.07	0.16	15.00
	35–54	6.71	5.52	4.64	0.26	17.13
	55+	8.73	2.84	2.17	0.05	13.80
	manjka	0.26	0.68	0.79	0.16	1.90
	skupaj	21.36	15.16	10.67	0.63	47.83
ženske	15–34	4.65	5.77	4.02	0.21	14.64
	35–54	6.67	5.35	4.53	0.11	16.65
	55+	12.79	2.94	1.67	0.16	17.56
	manjka	0.47	0.53	0.74	0.05	1.79
	skupaj	24.58	14.58	10.96	0.53	50.65
neznano	15–34	0.00	0.16	0.11	0.00	0.26
	35–54	0.05	0.00	0.05	0.00	0.11
	55+	0.21	0.05	0.00	0.00	0.26
	manjka	0.26	0.11	0.11	0.42	0.90
	skupaj	0.53	0.32	0.26	0.42	1.53
skupaj		46.47	30.06	21.89	1.58	100.00

Tako dobljeno tabelo množimo s številom prebivalcev v populacijski tabeli in dobimo končno tabelo, primerno za uteževanje.

**Tabela 6.2.1.6: Prirejena populacija spol/starost/izobrazba (N), Belgija 2002/R1**

spol	starost	izobrazba				skupaj
		1 nizka	2 srednja	3 visoka	0 manjkajoča	
moški	15-34	480454.3	520965.2	260796.1	13435.2	1275650.7
	35-54	570791.7	469199.6	394788.3	22392.0	1457171.6
	55+	742826.0	241231.6	184749.9	4478.4	1173285.9
	manjka	22392.0	58219.2	67176.0	13435.2	161222.4
	skupaj	1816464.0	1289615.5	907510.3	53740.8	4067330.6
ženske	15-34	395040.3	490453.6	342059.0	17913.6	1245466.6
	35-54	567458.2	455027.9	384838.2	8956.8	1416281.2
	55+	1087651.8	249889.5	142287.9	13435.2	1493264.4
	manjka	40305.6	44784.0	62697.6	4478.4	152265.6
	skupaj	2090455.9	1240155.1	931882.8	44784.0	4307277.8
neznano	15-34	0.0	13435.2	8956.8	0.0	22392.0
	35-54	4478.4	0.0	4478.4	0.0	8956.8
	55+	17913.6	4478.4	0.0	0.0	22392.0
	manjka	22392.0	8956.8	8956.8	35827.2	76132.8
	skupaj	44784.0	26870.4	22392.0	35827.2	129873.6
	skupaj	3951703.9	2556641.0	1861785.1	134352.0	8504482.0

Navajam še tabelo za uteževanje za regijo, kjer je bilo potrebno zamenjati prvi dve kategoriji, tako da so v enakem vrstnem redu kot na vzorcu:

**Tabela 6.2.1.7: Prirejena populacija regija (N), Belgija 2002/R1**

Regija	N
Flanders = Flemish	5017435.1
Brussels	817541.3
Wallonia	2762735.9
Skupaj	8597712.3

## 6.2.2 Belgija R3

Najprej pogledamo populacijski tabeli (Tabela 6.1.1.20, Tabela 6.1.1.21) in vzorec (Tabela 6.2.2.1 in Tabela 6.2.2.2):

**Tabela 6.2.2.1: Vzorec spol/starost/izobrazba (n), Belgija R3**

spol	starost	izobrazba				skupaj
		0 manjkajoča	1 nizka	2 srednja	3 visoka	
moški	15-34	0	80	116	66	262
	35-54	1	59	127	109	296
	55+	0	145	85	52	282
	skupaj	1	284	328	227	840
	ženske	0	90	111	75	276
ženske	35-54	0	81	135	145	361
	55+	0	190	78	53	321
	skupaj	0	361	324	273	958
	skupaj	1	645	652	500	1798

**Tabela 6.2.2.2: Vzorec regija (n), Belgija R3**

Regija	n
1 Flemish region	1128
2 Brussels region	99
3 Walloon region	571
Skupaj	1798

Kot lahko vidimo, vsebujeta obe tabeli za spol/starost/izobrazbo manjkajoče vrednosti, a na različnih mestih. Na populaciji manjka znana izobrazba za ženske od 15 do 34 let, v vzorcu pa za moške stare od 35 do 54 let. Potrebno je torej prilagajanje podatkov.

Podatke prilagodimo najprej na populaciji. Ker je odstotek manjkajočih vrednosti<sup>M</sup> za populacijo manjši od 1 % (0.002 %), te vrednosti prerazporedimo med ostale na sledeči način: izračunamo razmerje<sup>B</sup> med znanimi izobrazbenimi kategorijami<sup>A</sup> (predvidevamo, da se neznane vrednosti enakomerno porazdeljujejo), glede na razmerje razdelimo<sup>M\*B</sup> manjkajoče vrednosti<sup>C</sup> in seštejemo<sup>A + C</sup>. Vsota nam da prerazporejene vrednosti<sup>D</sup>.

**Tabela 6.2.2.3: Prerazporejanje manjkajočih vrednosti na populaciji, Belgija 2006**

	1 nizka	2 srednja	3 visoka	manjkajoča	Skupaj
vse ženske 15–35	389190.6	515493.5	403963.4	207.0 <sup>M</sup>	1308854.5
z znano izobrazbo <sup>A</sup>	389190.6	515493.5	403963.4		1308647.5
% znane izobrazbe <sup>B</sup>	29.74%	39.39%	30.87%		
razporejene manjkajoče <sup>C</sup>	61.6	81.5	63.9		
prerazporejene vrednosti <sup>D</sup>	389252.2	515575.0	404027.3		1308854.5

Tako dobimo novo populacijsko tabelo brez manjkajočih vrednosti:

**Tabela 6.2.2.4: Populacija spol/starost/izobrazba brez manjkajočih vrednosti (N), Belgija 2006**

spol	starost	izobrazba			skupaj
		nizka	srednja	visoka	
moški	15–34	466678.4	563880.1	304101.1	1334659.7
	35–54	524903.2	564633.7	469634.3	1559171.2
	55+	768233.7	326070.2	255706.6	1350010.5
	skupaj	1759815.3	1454584.0	1029442.0	4243841.3
ženske	15–34	389252.2	515575.0	404027.3	1308854.5
	35–54	491996.4	554834.8	485418.7	1532249.9
	55+	1123096.3	325687.3	215631.2	1664414.8
	skupaj	2004345.0	1396097.1	1105077.2	4505519.3
	skupaj	3764160.2	2850681.1	2134519.3	8749360.6

Sedaj se lotimo vzorca. Ker imamo manjkajoče vrednosti samo za eno spremenljivko, za izobrazbo, izračunamo odstotke, tako da razmerja za ostali dve spremenljivki ostanetjo enaka, izračunamo torej samo odstotke po vrsticah.

**Tabela 6.2.2.5: Vzorec spol/starost/izobrazba (%), Belgija R3**

spol	starost	izobrazba				skupaj
		0 manjkajoča	1 nizka	2 srednja	3 visoka	
moški	15–34	0.00	30.53	44.27	25.19	100
	35–54	0.34	19.93	42.91	36.82	100
	55+	0.00	51.42	30.14	18.44	100
	skupaj	0.34	101.89	117.32	80.45	300
ženske	15–34	0.00	32.61	40.22	27.17	100
	35–54	0.00	22.44	37.40	40.17	100
	55+	0.00	59.19	24.30	16.51	100
	skupaj	0.00	114.24	101.91	83.85	300
skupaj		0.34	216.12	219.23	164.31	600

Isto naredimo za populacijo:

**Tabela 6.2.2.6: Populacija spol/starost/izobrazba brez manjkajočih vrednosti (%), Belgija 2006**

spol	starost	izobrazba			skupaj
		1 nizka	2 srednja	3 visoka	
moški	15–34	34.97	42.25	22.78	100.00
	35–54	33.67	36.21	30.12	100.00
	55+	56.91	24.15	18.94	100.00
	skupaj	125.54	102.62	71.85	300.00
ženske	15–34	29.74	39.39	30.87	100.00
	35–54	32.11	36.21	31.68	100.00
	55+	67.48	19.57	12.96	100.00
	skupaj	129.33	95.17	75.50	300.00
skupaj		254.86	197.79	147.35	600.00

Vstavimo manjkajoče vrednosti iz vzorca in preračunamo populacijske odstotke za moške, 35–54 let, ker se le tu pojavi manjkajoča vrednost: izračun za nižjo izobrazbo =  $33.67 * (100 \% - 0.34 \%) = 33.55$ .

**Tabela 6.2.2.7: Dopolnjena populacija spol/starost/izobrazba (%), Belgija 2006/R3**

spol	starost	izobrazba				skupaj
		nizka	srednja	visoka	manjkajoča	
moški	15–34	34.97	42.25	22.78	0.00	100.00
	35–54	33.55	36.09	30.02	0.34	100.00
	55+	56.91	24.15	18.94	0.00	100.00
	skupaj	125.42	102.49	71.75	0.34	300.00
ženske	15–34	29.74	39.39	30.87	0.00	100.00
	35–54	32.11	36.21	31.68	0.00	100.00
	55+	67.48	19.57	12.96	0.00	100.00
	skupaj	129.33	95.17	75.50	0.00	300.00
skupaj		254.75	197.66	147.25	0.34	600.00

Tabela 6.2.2.7 množimo z N in dobimo končno prirejeno populacijsko tabelo za spol/starost/izobrazbo. Regijo se obravnava kot pri prejšnjem primeru.

**Tabela 6.2.2.8: Prirejena populacija spol/starost/izobrazba (N), Belgija 2006/R3**

spol	starost	izobrazba				skupaj
		nizka	srednja	visoka	manjkajoča	
moški	15–34	466678.4	563880.1	304101.1	0.0	1334659.7
	35–54	523129.8	562726.2	468047.7	5267.5	1559171.2
	55+	768233.7	326070.2	255706.6	0.0	1350010.5
	skupaj	1758042.0	1452676.4	1027855.4	5267.5	4243841.3
ženske	15–34	389252.2	515575.0	404027.3	0.0	1308854.5
	35–54	491996.4	554834.8	485418.7	0.0	1532249.9
	55+	1123096.3	325687.3	215631.2	0.0	1664414.8
	skupaj	2004345.0	1396097.1	1105077.2	0.0	4505519.3
skupaj		3762386.9	2848773.5	2132932.7	5267.5	8749360.6

## 6.2.3 Slovenija R2

Pogledamo populacijski tabeli (Tabela 6.1.2.6 in Tabela 6.1.2.7) ter vzorčni tabeli (Tabela 6.2.3.1 in Tabela 6.2.3.2). Na populaciji imamo manjkajoče vrednosti za izobrazbo (0.053 %), na vzorcu pa manjkajoče vrednosti za spol, starost in izobrazbo.

**Tabela 6.2.3.1: Vzorec spol/starost/izobrazba (n), Slovenija R2**

spol	starost	izobrazba				skupaj
		0 manjkajoča	1 nizka	2 srednja	3 visoka	
neznano	0	1	1	3	0	5
	15–34	0	2	1	2	5
	35–54	0	3	7	4	14
	55+	0	5	2	1	8
	skupaj	1	11	13	7	32
moški	0	0	1	0	0	1
	15–34	0	128	80	22	230
	35–54	0	111	86	30	227
	55+	0	121	52	17	190
	skupaj	0	361	218	69	648
ženske	0	0	4	3	1	8
	15–34	1	90	110	30	231
	35–54	1	112	105	32	250
	55+	0	206	57	10	273
	skupaj	2	412	275	73	762
skupaj		3	784	506	149	1442

**Tabela 6.2.3.2: Vzorec regija (n), Slovenija R2**

Regija	n
1 Vzhodna Slovenija	826
2 Zahodna Slovenija	616
Skupaj	1442

Najprej prerazporedimo manjkajoče vrednosti med populacijo (kot smo naredili pri poglavju Belgija R3 – Tabela 6.2.2.3), le da tu to naredimo za moške in ženske 15–34 let), nato sledi postopek vstavljanja manjkajočih vrednosti, podobno kot pri Belgiji R1. Regija je različna, na vzorcu so kategorije NUTS 3, za populacijo kategorije NUTS 2, zato združimo vzorčne kategorije v NUTS 2.

**Tabela 6.2.3.3: Prirejena populacija spol/starost/izobrazba (N), Slovenija 2004/R2**

spol	starost	izobrazba				skupaj
		1 nizka	2 srednja	3 visoka	0 manjkajoča	
moški	15–34	80351.4	179860.2	25948.2	0.0	286159.7
	35–54	57374.5	197623.9	47120.3	0.0	302118.7
	55+	56813.9	122586.5	35634.5	0.0	215034.9
	manjka	1183.3	0.0	0.0	0.0	1183.3
	skupaj	195723.1	500070.5	108702.9	0.0	804496.6
ženske	15–34	68535.0	154014.6	48246.0	1183.3	271978.9
	35–54	72036.0	163966.3	56334.3	1183.3	293519.9
	55+	159068.4	103975.8	26009.2	0.0	289053.5
	manjka	4733.4	3550.0	1183.3	0.0	9466.8
	skupaj	304372.8	425506.8	131772.8	2366.7	864019.0
neznano	15–34	2366.7	1183.3	2366.7	0.0	5916.7
	35–54	3550.0	8283.4	4733.4	0.0	16566.8
	55+	5916.7	2366.7	1183.3	0.0	9466.8
	manjka	1183.3	3550.0	0.0	1183.3	5916.7
	skupaj	13016.8	15383.5	8283.4	1183.3	37867.0
	skupaj	513112.7	940960.8	248759.2	3550.0	1706382.7

## 6.2.4 Slovenija R3

Poglejmo si populacijski (Tabela 6.1.2.8, Tabela 6.1.2.9) in vzorčni tabeli (Tabela 6.2.4.1 in Tabela 6.2.4.2). Na populacijski tabeli nimamo manjkajočih vrednosti, imamo pa jih na vzorcu.

**Tabela 6.2.4.1: Vzorec spol/starost/izobrazba (n), Slovenija R3**

spol	starost	izobrazba				skupaj
		0 manjkajoča	1 nizka	2 srednja	3 visoka	
moški	0	0	0	1	2	3
	15–34	0	109	91	24	224
	35–54	0	110	81	26	217
	55+	0	123	76	24	223
	skupaj	0	342	249	76	667
ženske	0	0	2	0	0	2
	15–34	0	79	113	41	233
	35–54	0	122	88	62	272
	55+	2	205	70	25	302
	skupaj	2	408	271	128	809
	skupaj	2	750	520	204	1476

**Tabela 6.2.4.2: Vzorec regija (n), Slovenija R3**

Regija	n
1 Vzhodna Slovenija	837
2 Zahodna Slovenija	639
Skupaj	1476

Tokrat nam manjkajo vrednosti pri dveh spremenljivkah, starosti in izobrazbi. Vendar pri izobrazbi manjkajo vrednosti samo za ženski spol, tako da se lotimo računanja odstotkov posebej za moške in posebej za ženske. Za moške izračunamo odstotke po stolpcih, da ohranimo izobrazbena razmerja, ker nam manjka samo starost, za ženske pa ohranimo razmerja med izobrazbo in spolom:

**Tabela 6.2.4.3: Vzorec spol/starost/izobrazba (%), Slovenija R3**

spol	starost	izobrazba				skupaj
		0 manjkajoča	1 nizka	2 srednja	3 visoka	
moški	0	0	0.00	0.40	2.63	3.03
	15–34	0	31.87	36.55	31.58	100.00
	35–54	0	32.16	32.53	34.21	98.90
	55+	0	35.96	30.52	31.58	98.07
	skupaj	0	100.00	100.00	100.00	300.00
ženske	0	0.00	0.25	0.00	0.00	0.25
	15–34	0.00	9.77	13.97	5.07	28.80
	35–54	0.00	15.08	10.88	7.66	33.62
	55+	0.25	25.34	8.65	3.09	37.33
	skupaj	0.25	50.43	33.50	15.82	100.00
skupaj		0.25	150.43	133.50	115.82	400.00

Enako izračunamo odstotke za populacijo in dobimo naslednjo tabelo:

**Tabela 6.2.4.4: Populacija spol/starost/izobrazba (%), Slovenija 2006**

spol	starost	izobrazba			skupaj
		nizka	srednja	visoka	
moški	15–34	38.17	35.51	23.36	97.04
	35–54	28.13	39.32	44.86	112.32
	55+	33.70	25.18	31.77	90.65
	skupaj	100.00	100.00	100.00	300.00
ženske	15–34	7.01	17.15	6.09	30.25
	35–54	7.36	19.12	8.02	34.49
	55+	17.89	13.75	3.62	35.26
	skupaj	32.25	50.02	17.74	100.00
skupaj		132.25	150.02	117.74	400.00

Vstavimo odstotke v populacijsko tabelo in preračunamo populacijske odstotke.

Pri ženskah od celote (100 %) odštejemo vsoto vseh manjkajočih vrednosti.

Primer za moške: 15–34 let, višja izobrazba:  $23.36 * (100 \% - 2.63 \%) = 22.75$

Primer za ženske: 55+ let, nižja izobrazba:  $17.89 * (100 \% - (0.25 \% + 0.25 \%)) = 17.80$

**Tabela 6.2.4.5: Dopolnjena populacija spol/starost/izobrazba (%), Slovenija 2006/R3**

spol	starost	izobrazba				skupaj
		1 nizka	2 srednja	3 visoka	0 manjkajoča	
moški	15-34	38.17	35.36	22.75	0	96.28
	35-54	28.13	39.16	43.68	0	110.98
	55+	33.70	25.07	30.94	0	89.71
	manjka	0.00	0.40	2.63	0	3.03
	skupaj	100.00	100.00	100.00	0	300.00
ženske	15-34	6.97	17.07	6.06	0.00	30.10
	35-54	7.32	19.02	7.98	0.00	34.32
	55+	17.80	13.68	3.61	0.25	35.33
	manjka	0.25	0.00	0.00	0.00	0.25
	skupaj	32.34	49.77	17.65	0.25	100.00
skupaj		132.34	149.77	117.65	0.25	400.00

Množimo še z N in dobimo končno prirejeno populacijsko tabelo:

**Tabela 6.2.4.6: Prirejena populacija spol/starost/izobrazba (N), Slovenija 2006/R3**

spol	starost	izobrazba				skupaj
		1 nizka	2 srednja	3 visoka	0 manjkajoča	
moški	15-34	73121.98	183032.02	29412.763	0	285566.766
	35-54	53895.49	202693.19	56485.252	0	313073.925
	55+	64561.42	129784.27	40002.764	0	234348.453
	manjka	0	2078.6673	3402.7238	0	5481.39101
	skupaj	191578.9	517588.15	129303.5	0	838470.535
ženske	15-34	61706.98	151051.71	53661.091	0	266419.785
	35-54	64781.49	168372.39	70621.394	0	303775.27
	55+	157550.5	121092.34	31922.979	2188.226	312754.059
	manjka	2188.226	0	0	0	2188.22581
	skupaj	286227.2	440516.44	156205.46	2188.226	885137.34
skupaj		477806.1	958104.58	285508.97	2188.226	1723607.88

## 6.2.5 Švedska R3

Še zadnji primer situacije s prilagajanjem manjkajočih vrednosti. Poglejmo si populacijski tabeli (Tabela 6.1.3.5 in Tabela 6.1.3.6) in vzorčni tabeli (Tabela 6.2.5.1 in Tabela 6.2.5.2).

**Tabela 6.2.5.1: Vzorec spol/starost/izobrazba (n), Švedska R3**

spol	starost	izobrazba				skupaj
		0 manjkajoča	1 nizka	2 srednja	3 visoka	
neznano	0	1	0	0	0	1
	skupaj	1	0	0	0	1
moški	15-34	0	67	151	86	304
	35-54	1	57	136	115	309
	55+	2	152	95	89	338
	skupaj	3	276	382	290	951
	ženske	0	58	116	104	278
35-54		0	46	119	153	318
55+		1	222	61	95	379
skupaj		1	326	296	352	975
skupaj		5	602	678	642	1927

**Tabela 6.2.5.2: Vzorec regija (n), Švedska R3**

Regija	n
1 Stockholm	363
2 Östra Mellansverige	346
3 Sydsverige	281
4 Norra mellansveige	189
5 Mellemska Norrland	85
6 Övre Norrland	112
7 Småland med Öarna	176
8 Västsverige	375
Skupaj	1927

**Tabela 6.2.5.3: Vzorec spol/starost/izobrazba (%), Švedska R3**

spol	starost	izobrazba				skupaj
		0 manjkajoča	1 nizka	2 srednja	3 visoka	
neznano	0	0.05	0	0	0	0.05
	skupaj	0.05	0	0	0	0.05
moški	15-34	0.00	3.48	7.84	4.46	15.78
	35-54	0.05	2.96	7.06	5.97	16.04
	55+	0.10	7.89	4.93	4.62	17.54
	skupaj	0.16	14.32	19.82	15.05	49.35
ženske	15-34	0.00	3.01	6.02	5.40	14.43
	35-54	0.00	2.39	6.18	7.94	16.50
	55+	0.05	11.52	3.17	4.93	19.67
	skupaj	0.05	16.92	15.36	18.27	50.60
skupaj		0.26	31.24	35.18	33.32	100.00

Za razliko od prejšnjih situacij imamo tu več manjkajočih vrednosti na populaciji, za izobrazbo (13.93 %) in regijo (10.62 %). Na vzorcu pa za regijo nimamo manjkajočih vrednosti. Potrebno je obravnavati manjkajoče vrednosti v obeh populacijskih tabelah, na prejšnjih primerih je bilo to potrebno samo za tabelo spol/starost/izobrazba. Ker je manjkajočih vrednosti več kot 1 %, nam lahko njihovo prerazporejanje da slabše rezultate, ker več tvegamo. Zato utežujemo na dva načina: prvič, da vseeno prerazporedimo manjkajoče vrednosti med populacijo, in drugič, da ignoriramo manjkajoče vrednosti na populaciji ter uporabimo samo znane vrednosti. Na vzorcu ima regija vse znane vrednosti, pri tabeli spol/starost/izobrazba pa imamo povsod manjkajoče vrednosti.

**Način 1:**

Prvi način je podoben, kot so bili zgornji primeri. Manjkajoče vrednosti glede na razmerje med znanimi vrednostmi prerazporedimo med znane vrednosti. Nova tabela spol/starost/izobrazba je tako naslednja:

**Tabela 6.2.5.4: Populacija spol/starost/izobrazba (N), Švedska 2006, način 1**

spol	starost	izobrazba			skupaj
		nizka	srednja	visoka	
moški	15–34	320991.9	604041.4	241480.7	1166514.0
	35–54	197300.3	748639.9	296915.4	1242855.6
	55+	434699.3	581554.2	277893.0	1294146.5
	skupaj	952991.5	1934235.4	816289.2	3703516.1
ženske	15–34	280897.4	530853.4	311678.4	1123429.2
	35–54	134688.3	653111.5	415030.9	1202830.6
	55+	467358.3	657686.2	368578.9	1493623.5
	skupaj	882943.9	1841651.1	1095288.3	3819883.3
skupaj		1835935.5	3775886.5	1911577.4	7523399.4

Nova tabela za regijo pa naslednja:

**Tabela 6.2.5.5: Populacija regija (N), Švedska 2006, način 1**

Regija	N
1 Stockholm	1591526.3
2 Östra Mellansverige	1259146.1
4 Sydsverige	1099252.6
6 Norra Mellansverige	680045.4
7 Mellersta Norrland	305145.1
8 Övre Norrland	427338.6
9 Småland med öarna	656084.9
0A Västsverige	1504860.5
Skupaj	7523399.4

Odstotke spet izračunamo, tako da je vsota vseh 100 % (kot pri Belgiji R1 in Sloveniji R2). Vstavimo odstotke iz vzorca, preračunamo populacijske odstotke in dobimo končno prirejeno populacijsko tabelo:

**Tabela 6.2.5.6: Prirejena populacija spol/starost/izobrazba (N), Švedska 2006/R3, način 1**

spol	starost	izobrazba				skupaj
		1 nizka	2 srednja	3 visoka	manjkajoča	
moški	15–34	320159.0	602474.0	240854.2	0.0	1163487.2
	35–54	196788.4	746697.4	296145.0	3904.2	1243535.0
	55+	433571.4	580045.2	277172.0	7808.4	1298597.0
	skupaj	950518.8	1929216.6	814171.1	11712.6	3705619.2
ženske	15–34	280168.5	529476.0	310869.7	0.0	1120514.3
	35–54	134338.8	651416.8	413954.0	0.0	1199709.7
	55+	466145.6	655979.7	367622.6	3904.2	1493652.1
	skupaj	880652.9	1836872.6	1092446.3	3904.2	3813876.1
neznano	0	0.0	0.0	0.0	3904.2	3904.2
	skupaj	0.0	0.0	0.0	3904.2	3904.2
	skupaj	1831171.8	3766089.2	1906617.4	19521.0	7523399.4

## Način 2:

Pri tem načinu iz populacijske tabele odstranimo manjkajoče vrednosti. Ker ima regija manjši odstotek manjkajočih vrednosti, se najprej izbrisati manjkajočih vrednosti za izobrazbo. V tabeli spol/starost/izobrazba se številke znotraj celic ne spremenijo, pač pa se spremeni vsota, število N je manjše. Ker je regija ločena od tabele spol/starost/izobrazba, se vsak posameznik pojavi v obeh tabelah. Obstaja možnost, da za nekoga nimamo znane izobrazbe, pač pa imamo znano regijo. Zato moramo iz 4D-tabele (Tabela B3.3: Švedska 2006) izbrisati manjkajočo izobrazbo in ponovno izračunati vrednosti za regijo. Manjkajoče vrednosti za regijo tu pokrivajo del manjkajočih vrednosti za izobrazbo<sup>5</sup>. Novi populacijski tabeli sta naslednji:

**Tabela 6.2.5.7: Populacija spol/starost/izobrazba (N), Švedska R3, način 2**

spol	starost	izobrazba			skupaj
		nizka	srednja	visoka	
moški	15–34	289707.6	545170.7	217945.7	1052824.0
	35–54	196458.8	745446.6	295648.9	1237554.2
	55+	326210.8	436414.8	208538.9	971164.5
	skupaj	812377.1	1727032.1	722133.5	3261542.7
ženske	15–34	254567.8	481094.5	282463.6	1018125.9
	35–54	134117.4	650343.2	413271.8	1197732.4
	55+	312221.9	439371.8	246231.7	997825.4
	skupaj	700907.1	1570809.5	941967.1	3213683.7
skupaj		1513284.2	3297841.7	1664100.6	6475226.4

**Tabela 6.2.5.8: Populacija regija (N), Švedska R3, način 2**

Regija	N
1 Stockholm	1359564.7
2 Östra Mellansverige	1086782.5
4 Sydsverige	937017.8
6 Norra Mellansverige	585847.09
7 Mellersta Norrland	264995.1
8 Övre Norrland	372504.76
9 Småland med öarna	568021.5
0A Västsverige	1300493
Skupaj	6475226.4

<sup>5</sup> Če pa bi nato še vedno ostale manjkajoče vrednosti za regijo, iz 4D-tabele odstranimo še te in nato ponovno izračunamo vrednosti za spol/starost/izobrazbo in za regijo.

Vzorčni odstotki so enaki kot v prvem načinu. Spet sledi postopek vstavljanja manjkajočih vrednosti in preračunavanje ostalih. Za regijo je končna tabela za uteževanje zgornja, Tabela 6.2.5.8, za spol/starost/izobrazba pa naslednja:

**Tabela 6.2.5.9: Prijeljena populacija spol/starost/izobrazba (N), Švedska 2006/R3, način 2**

spol	starost	izobrazba				skupaj
		1 nizka	2 srednja	3 visoka	manjkajoča	
moški	15–34	288955.9	543756.1	217380.2	0.0	1050092.2
	35–54	195949.0	743512.4	294881.8	3360.3	1237703.4
	55+	325364.4	435282.4	207997.8	6720.5	975365.1
	skupaj	810269.2	1722551.0	720259.7	10080.8	3263160.7
ženske	15–34	253907.3	479846.2	281730.7	0.0	1015484.1
	35–54	133769.4	648655.8	412199.5	0.0	1194624.7
	55+	311411.8	438231.8	245592.8	3360.3	998596.6
	skupaj	699088.4	1566733.7	939523.0	3360.3	3208705.4
neznano	0	0.0	0.0	0.0	3360.3	3360.3
	skupaj	0.0	0.0	0.0	3360.3	3360.3
skupaj		1509357.7	3289284.7	1659782.7	16801.3	6475226.4

### 6.3 Potek uteževanja in analiza uteženih podatkov

Uteževanje povsod poteka po naslednjih korakih: najprej ustvarim začetno utež, spremenljivko  $w_{psN}$ , ki postane  $dweight$ . Nato jo množim s številom populacije deljeno s številom enot v vzorcu. Sledi postopek uteževanja (v tem primeru raking na dve tabeli, saj imamo za poststratifikacijo premajhen vzorec), kjer najprej vzorec utežimo z  $w_{psN}$  in nato izvajamo postopek rakinga tako dolgo, dokler ni utežen vzorec enak prirejenim populacijskim tabelam. Nato ustvarim novo spremenljivko,  $w_{ps.n}$ , kjer  $w_{psN}$  množim s številom enot v vzorcu in delim s številom populacije. Ko vzorec utežim s to utežjo, dobim takšno število enot, kot bi bilo, če bi utežila samo z  $dweight$ . Razmerja ostanejo kot v populaciji, aritmetična sredina uteži pa je 1, tako da lahko izračunam VIF in ugotovim, kako se je povečala varianca zaradi uteževanja.

Najprej naj predstavim pregled odstotkov osnovne in spremenjene populacije ter neuteženega in uteženega vzorca za vse tri države. Kot lahko vidimo, se populacijska tabela prilagodi odstotku neznanih v vzorcu. Razlika je edino za

Slovenijo v R2, saj smo odstotke prilagodili različno glede na spol. Dodana je tudi vrstica za utežen vzorec brez manjkajočih vrednosti.

**Tabela 6.2.5.1: Primerjava populacije in vzorca (%), Belgija**

BE	meritev	spol			starost				izobrazba				
		vir/vzorec	m	ž	nezn.	15–34	35–54	55+	nezn.	1	2	3	neznano
R1 2002	LFS	48.44	51.56	/		31.20	35.61	33.19	/	48.17	30.41	21.42	/
	neutežen	50.71	47.76	1.53		31.33	35.91	28.17	4.58	34.49	36.81	27.12	1.58
	LFS*	47.83	50.65	1.53		29.91	33.89	31.62	4.58	46.47	30.06	21.89	1.58
	utežen	47.83	50.65	1.53		29.91	33.89	31.62	4.58	46.47	30.06	21.89	1.58
	utežen <sup>BM</sup>	48.57	51.43	/		31.34	35.52	33.14	/	47.21	30.54	22.24	/
R2 2004	LFS	48.48	51.52	/		30.70	35.50	33.81	/	45.99	31.48	22.53	/
	neutežen	49.16	50.84	/		31.72	36.84	31.44	/	33.86	38.64	27.45	0.06
	LFS*	48.48	51.52	/		30.70	35.50	33.81	/	45.97	31.46	22.51	0.05
	utežen	48.48	51.52	/		30.70	35.50	33.81	/	45.97	31.46	22.51	0.05
	utežen <sup>BM</sup>	48.48	51.52	/		30.70	35.50	33.81	/	45.99	31.48	22.53	/
R3 2006	LFS	48.50	51.50	/		30.21	35.33	34.45	/	43.02	32.58	24.40	0.00
	neutežen	46.72	53.28	/		29.92	36.54	33.54	/	35.87	36.26	27.81	0.06
	LFS*	48.50	51.50	/		30.21	35.33	34.45	/	43.00	32.56	24.38	0.06
	utežen	48.50	51.50	/		30.21	35.33	34.45	/	43.00	32.56	24.38	0.06
	utežen <sup>BM</sup>	48.50	51.50	/		30.21	35.33	34.45	/	43.03	32.58	24.39	/

Opombe: \* prirejena populacijska tabela; <sup>BM</sup> samo znane vrednosti, brez manjkajočih

**Tabela 6.2.5.2: Primerjava populacije in vzorca (%), Slovenija**

SI	meritev	spol			starost				izobrazba				
		vir	m	ž	nezn.	15–34	35–54	55+	nezn.	1	2	3	neznano
R1 2002	LFS	48.49	51.51	/		34.20	35.95	29.85	/	31.67	55.84	10.85	1.64
	neutežen	47.60	52.34	0.07		33.44	37.00	29.56	/	53.72	31.01	14.68	0.59
	LFS <sup>*</sup> 1	48.44	51.49	0.07		34.20	35.95	29.85	/	31.97	56.44	10.99	0.60
	LFS <sup>*</sup> 2	48.51	51.42	0.07		34.18	35.85	29.97	/	32.02	56.42	10.97	0.60
	utežen1	48.44	51.49	0.07		34.20	35.95	29.85	/	31.97	56.44	10.99	0.60
	utežen2	48.51	51.42	0.07		34.18	35.85	29.97	/	32.02	56.42	10.97	0.60
	utežen <sup>BM1</sup>	48.47	51.53	/		34.22	35.97	29.80	/	31.97	88.42	99.40	/
	utežen <sup>BM2</sup>	48.54	51.46	/		34.20	35.88	29.92	/	32.21	56.75	11.03	/
R2 2004	LFS	48.52	51.48	/		33.64	35.91	30.45	/	29.84	55.66	14.45	0.05
	neutežen	44.94	52.84	2.22		32.32	34.05	32.66	0.97	54.37	30.10	15.33	0.21
	LFS*	47.15	50.63	2.22		33.06	35.88	30.10	0.97	30.07	55.14	14.58	0.21
	utežen	47.15	50.63	2.22		33.06	35.88	30.10	0.97	30.07	55.14	14.58	0.21
	utežen <sup>BM</sup>	48.22	51.78	/		33.38	36.23	30.39	/	30.13	55.26	14.61	/
R3 2006	LFS	48.65	51.35	/		32.19	36.01	31.80	/	27.68	55.71	16.61	/
	neutežen	45.19	54.81	/		30.96	33.13	35.57	0.34	50.81	35.23	13.82	0.14
	LFS*	48.65	51.35	/		32.03	35.79	31.74	0.44	27.72	55.59	16.56	0.13
	utežen	48.65	51.35	/		32.03	35.79	31.74	0.44	27.72	55.59	16.56	0.13
	utežen <sup>BM</sup>	48.65	51.35	/		32.17	35.95	31.88	/	27.76	55.66	16.59	/

Opombe: \* prirejena populacijska tabela; <sup>1</sup> na prvi način; <sup>2</sup> na drugi način; <sup>BM</sup> samo znane vrednosti, brez manjkajočih (številka zraven označuje način)

**Tabela 6.2.5.3: Primerjava populacije in vzorca (%), Švedska**

SE	meritev	spol			starost				izobrazba				
		vir	m	ž	nezn.	15–34	35–54	55+	nezn.	1	2	3	neznano
R1 2002	LFS	49.10	50.90	/		30.54	33.47	35.99	/	21.94	44.96	19.19	13.91
	neutežen	50.73	49.07	0.20		29.96	34.97	34.87	0.20	47.67	21.51	30.57	0.25
	LFS <sup>*1</sup>	49.02	50.78	0.20		30.46	33.39	35.95	0.20	26.83	50.97	21.95	0.25
	LFS <sup>*2</sup>	50.32	49.48	0.20		32.34	38.52	28.94	0.20	25.42	52.10	22.23	0.25
	utežen1	49.02	50.78	0.20		30.46	33.39	35.95	0.20	26.83	50.97	21.95	0.25
	utežen2	50.32	49.48	0.20		32.34	38.52	28.94	0.20	25.42	52.10	22.23	0.25
	utežen <sup>BM1</sup>	49.12	50.88	/		30.52	33.46	36.02	/	26.89	51.10	22.01	/
	utežen <sup>BM2</sup>	50.42	49.58	/		32.41	38.60	28.99	/	25.48	52.23	22.29	/
R2 2004	LFS	49.17	50.83	/		30.34	33.03	36.63	/	21.11	44.59	20.36	13.94
	neutežen	50.36	49.64	/		29.41	34.50	36.09	/	47.33	22.38	29.93	0.36
	LFS <sup>*1</sup>	49.17	50.83	/		30.34	33.03	36.63	/	25.70	50.70	23.24	0.15
	LFS <sup>*2</sup>	50.36	49.64	/		31.96	38.16	29.88	/	24.44	51.63	23.57	0.16
	utežen1	49.17	50.83	/		30.34	33.03	36.63	/	25.70	50.70	23.24	0.15
	utežen2	50.36	49.64	/		31.96	38.16	29.88	/	24.44	51.63	23.57	0.16
	utežen <sup>BM1</sup>	49.17	50.83	/		30.34	33.03	36.63	/	25.79	50.88	23.33	/
	utežen <sup>BM2</sup>	50.36	49.64	/		31.96	38.16	29.88	/	24.53	51.81	23.66	/
R3 2006	LFS	49.23	50.77	/		30.44	32.51	37.05	/	20.11	43.83	22.12	13.93
	neutežen	49.35	50.60	0.05		30.20	32.54	37.21	0.05	33.32	35.18	31.24	0.26
	LFS <sup>*1</sup>	49.25	50.69	0.05		30.36	32.48	37.11	0.05	24.34	50.06	25.34	0.26
	LFS <sup>*2</sup>	50.39	49.55	0.05		31.90	37.56	30.48	0.05	23.31	50.80	25.63	0.26
	utežen1	49.25	50.69	0.05		30.36	32.48	37.11	0.05	24.34	50.06	25.34	0.26
	utežen2	50.39	49.55	0.05		31.90	37.56	30.48	0.05	23.31	50.80	25.63	0.26
	utežen <sup>BM1</sup>	49.28	50.72	/		30.37	32.49	37.13	/	24.40	50.19	25.41	/
	utežen <sup>BM2</sup>	50.42	49.58	/		31.92	37.58	30.50	/	23.37	50.93	25.70	/

Opombe: \* prirejena populacijska tabela; <sup>\*1</sup> na prvi način; <sup>\*2</sup> na drugi način; <sup>BM</sup> samo znane vrednosti, brez manjkajočih (številka zraven označuje način)

Če za meritev pogledamo prvo vrstico (LFS) in zadnjo vrstico (utežen<sup>BM</sup>), naj bi bila razmerja enaka, kar drži tam, kjer so bile manjkajoče vrednosti samo na eni spremenljivki. Kjer pa so manjkajoče vrednosti na več kot eni spremenljivki, se tudi razmerja spremenijo, saj ima neka enota lahko manjkajočo vrednost za obe spremenljivki in tako vpliva na oba odstotka. Če bi sestavili kontingenčno tabelo (*angl. crosstabs*) med tistimi spremenljivkami, ki imajo manjkajoče vrednosti, ne da bi upoštevali te manjkajoče vrednosti, bi dobili enaka razmerja, kot so na populaciji med spremenljivkami.

V primeru, da bi manjkajoče vrednosti na vzorcu ignorirali, nam raking ne bi konvergiral, saj ima enota manjkajočo vrednost za eno spremenljivko, za ostale pa ne, in se tako upošteva samo v tisti tabeli za raking, kjer nima manjkajoče vrednosti, ne pa tudi v drugi.

Zanima nas še, kako je s povečanjem vzorčne variance. Sledi prikaz izračunanih opisnih statistik za *dweight* (ki je v tu enaka za vse meritve in vse države) in končno utež (*w\_ps.n* ali *w\_ps.n1*, *w\_ps.n2*) za vse tri meritve za izbrane države:

**Tabela 6.2.5.4: Pregled opisnih statistik končnih uteži za izbrane tri države**

Država	Meritve	ime uteži	n	minimum	maksimum	povprečje	std. napaka	std. odklon	varianca	VIF
BE, SI, SE	vse	dweight	/	1	1	1	0	0	0	1
BE	R1	w_ps.n	1899	0.55	3.48	1.0000	0.00881	0.38401	0.147	1.15
BE	R2	w_ps.n	1778	0.63	1.75	1.0000	0.00651	0.27450	0.075	1.08
BE	R3	w_ps.n	1798	0.61	3.50	1.0000	0.00700	0.29684	0.088	1.09
SI	R1	w_ps.n1	1519	0.33	2.11	1.0000	0.01293	0.50402	0.254	1.25
		w_ps.n2	1519	0.33	2.10	1.0000	0.01293	0.50393	0.254	1.25
SI	R2	w_ps.n	1442	0.39	2.22	1.0000	0.01439	0.54644	0.299	1.30
SI	R3	w_ps.n	1476	0.42	2.16	1.0000	0.01377	0.52889	0.280	1.28
SE	R1	w_ps.n1	1999	0.34	6.90	1.0000	0.02079	0.92975	0.864	1.86
		w_ps.n2	1999	0.37	5.24	1.0000	0.01968	0.87999	0.774	1.77
SE	R2	w_ps.n1	1948	0.23	9.35	1.0000	0.02195	0.96863	0.938	1.94
		w_ps.n2	1948	0.28	6.92	1.0000	0.02005	0.88513	0.783	1.78
SE	R3	w_ps.n1	1927	0.48	3.10	1.0000	0.01026	0.45043	0.203	1.20
		w_ps.n2	1927	0.38	2.38	1.0000	0.00975	0.42803	0.183	1.18

Kjer sem izračunala uteži na dva načina, vidim, da je boljši 2. način, saj je VIF-faktor pri tem manjši. Za Slovenijo v prvi meritvi ni razlike v VIF-faktorju, je pa tam odstotek manjkajočih vrednosti majhen (1.64 %), kar pomeni, da majhen odstotek manjkajočih vrednosti, ki jih prerazporedimo, nima velikega vpliva na celoten VIF. Pojavi se nam torej učinek povečanja vzorčne variance.

Če primerjam z odstotki manjkajočih vrednosti pri posameznih spremenljivkah, lahko opazim, da se VIF-faktor malo poveča pri tistih državah, kjer je odstotek manjši. Najmanj se poveča v primeru Belgije v prvih dveh meritvah, odstotek manjkajočih vrednosti je tam minimalen. Kjer je odstotek v višini do 5 %, je VIF-faktor večji, največji pa je v primeru Švedske, ki ima manjkajoče vrednosti na populaciji. Slednja ima višji VIF v prvih dveh meritvah, v tretji pa ne. Do tega verjetno pride, ker je izobrazbena struktura v neutreženem vzorcu v prvih dveh meritvah precej bolj različna od populacijske strukture, zato imajo določene enote večjo utež.

Višji VIF pomeni večjo mersko napako zaradi uteževanja. Kjer imamo višji odstotek manjkajočih vrednosti (na populaciji ali na vzorcu), je večja verjetnost, da nam uteževanje ne izboljša rezultatov, pač pa še poveča napake. V takšnih primerih je potrebno presoditi glede na dano situacijo, ali je uteževanje res smiselno ali je boljše ponoviti raziskavo. Če se vzorec namreč radikalno razlikuje od populacijske strukture (npr. Švedska v prvi in drugi meritvi) je razpon velikosti uteži precejšnji, kar je za določene spremenljivke lahko problematično.

## 7 Zaključek

Družboslovno raziskovanje je čedalje bolj razširjeno, povečuje se obseg raziskav in njihova pomembnost. Ena izmed večjih raziskav je tudi mednarodna raziskava, imenovana Evropska družboslovna raziskava (ESS), ki zajema več kot 20 evropskih držav in poteka vsaki dve leti, do sedaj so bile opravljene tri meritve. Raziskava vsebuje natančne verjetnostne vzorce in strogo metodologijo, a se kljub vsemu ne more izogniti problemu nepokritja in neodgovorov. Do njiju prihaja, ker je kljub enostavni definiciji populacije (vsi prebivalci države, starejši od 15 let, ki živijo v zasebnih gospodinjstvih) v nekaterih državah to težko aplicirati na realnost, države imajo starostne omejitve, dostop do določenih območij je lahko predrag ali prenevaren.

Neodgovori lahko vplivajo na kvaliteto anketnih statistik, saj imajo lahko nerespondenti drugačne vrednosti od vrednosti celotnega vzorca. V ESS želijo doseči vsaj 70 % stopnjo odgovorov, saj vedo, da se bodo vedno našli nerespondenti, ki jih anketarji ne bodo mogli locirati (nekontakti), ki ne bodo hoteli odgovarjati (zavnitve) ali ne bodo zmožni sodelovati. Poznamo različne pristope k zmanjšanju napake zaradi neodgovorov, med njimi je tudi uteževanje. Z uteževanjem lahko odstranimo pristranskost zaradi neodgovorov, vendar prihaja do povečanja vzorčne variance.

Poznamo dve osnovni metodi populacijskega uteževanja, in sicer poststratifikacijo (kjer imamo vse podatke v eni tabeli) ter raking (imamo podatke v več tabelah). V prvih dveh meritvah so uteževanje izvajali na tri kontrolne spremenljivke, spol, starost in izobrazbo, v tej diplomski nalogi pa je predstavljeno še za regijo. Za vsako kontrolno spremenljivko obstaja tudi mednarodna klasifikacija, po kateri so združene kategorije za uteževanje. V diplomskem delu sem se osredotočila na tri države, ki so približno dosegle želeno stopnjo odgovorov, a se vsaka sooča z drugačnimi problemi glede populacijskih podatkov. Te države so Belgija, Slovenija in Švedska.

Pred začetkom uteževanja je potrebno pridobiti primerne populacijske podatke za kontrolne spremenljivke, kar je lahko problematično. Obstaja namreč več virov, izmed katerih sem primerjala dva: ESS-prilogo 'populacijske statistike in druga dokumentacija' ter Eurostatovo Anketo o delovni sili (LFS). Preverjala sem, kateri vir populacijskih podatkov je primernejši za uteževanje, in ugotovila, da so primernejši LFS-jevi populacijski podatki, saj so za razliko od ESS-podatkov bolj ažurni ter bolj podrobni. LFS-podatke zbirajo na enem mestu in so tako primerljivi med državami, medtem ko ESS-podatke priskrbijo nacionalni koordinatorji vsak za svojo državo.

Podatke je bilo potrebno pridobiti na tak način, da jih lahko uredimo v ustrezne tabele za uteževanje glede na uporabljeni metodo. Najboljša metoda je sicer poststratifikacija, ker nam največ pove o enotah, vendar bi hitro prišlo do premajhnega števila enot v celicah za uteževanje, tako da sem za uteževanje kasneje uporabila raking metodo na tabeli spol/starost/izobrazba in regija.

Poleg neodgovora enot (nesodelovanje v raziskavi) poznamo še neodgovore spremenljivk (neodgovor na vprašanje). Pojavijo se lahko tako na populaciji kot na vzorcu in jih je potrebno primerno obravnavati. V drugem poglavju empirične analize predstavim situacije, na katere lahko naletimo ob manjkajočih vrednostih. Podatki lahko manjkajo za eno ali več spremenljivk, na vzorcu ali na populaciji ali celo na obeh. Obravnava teh manjkajočih podatkov je odvisna od števila spremenljivk, ki imajo manjkajoče vrednosti, od tega, kje manjkajo podatki, in od ujemanja med manjkajočimi vrednostmi med populacijo in vzorcem. Manjkajoče podatke na populaciji lahko prerazporedimo ali jih ignoriramo, tiste na vzorcu pa vstavimo v populacijsko tabelo in ostale vrednosti preračunamo, tako da utežen vzorec upošteva vse manjkajoče vrednosti. V primeru, da manjkajoče vrednosti na vzorcu ignoriramo, nam raking ne konvergira, saj ima enota lahko manjkajočo vrednost samo za eno spremenljivko, za ostale pa ne.

V prirejenih populacijskih tabelah se odstotki za posamezne kategorije spremenljivk razlikujejo od prvotnih, razlike pa so odvisne od števila spremenljivk znotraj tabele, kjer se manjkajoče vrednosti ponavljajo. Primer tabele spol/starost/izobrazba: če imamo manjkajoče vrednosti za eno spremenljivko (npr. izobrazbo), ostanejo odstotki za kategorije drugih dve (spol in starost) enaki, tista ena pa se spremeni. Če imamo manjkajoče vrednosti na dveh spremenljivkah, ostanejo razmerja enaka samo med kategorijami tretje spremenljivke.

Na podlagi uteži, pridobljenih z metodo rakinga, sem izračunala faktor povečanja variance za končno utež (VIF). Kjer je odstotek manjkajočih vrednosti večji, ne glede na to, ali se le-te pojavijo na populaciji ali na vzorcu, je tudi VIF-faktor večji. Izkaže se sicer, da je pri večjem odstotku populacijskih manjkajočih vrednosti le-te bolje ignorirati, vendar to lahko privede do spremembe razmerja med vrednostmi spremenljivk na populaciji, kar nam da drugačno porazdelitev prebivalstva, kot je v resnici. Uteževanje v takšnem primeru ni smiselno, saj obstaja večja verjetnost napake. Vendar je vsakokrat potrebno presoditi glede na trenutno situacijo odstopanja vzorca od populacije in glede na ciljne spremenljivke, ki jih želimo proučevati.

V prihodnje bi bilo potrebno raziskovanje razširiti še na specifične spremenljivke in za njih izračunati srednjo kvadratno napako. Le tako bi namreč dobili pravi vpogled v to, kako nam je uteževanje sploh koristilo.

## 8 Viri in literatura

Billiet, Jaak, Achim Koch in Michel Philippens (2007): Understanding and improving response rates. V Roger Jowell, Caroline Roberts, Rory Fitzgerald in Gillian Eva (ur.): *Measuring attitudes cross-nationally: lessons from the European Social Survey*, 113–137. Los Angeles: Sage Publications.

ESS (2008): *Weighting*. Dostopno na <http://ess-xml.nsd.uib.no/2.3.1.12.cfm> (10. avgust 2008).

ESS DATA (2003a): *ESS Round 1, Data download: ESS 1 – integrated file, edition 6.0*. Dostopno na <http://ess.nsd.uib.no/index.jsp?module=download&year=2003&country=&download=%5CDirect+Data+download%5C2003%5C01%23ESS1+-+integrated+file%2C+edition+6.0%5C.%5CESS1e06.spss.zip> (4. april 2008).

ESS DATA (2003b): *ESS Round 1, Deviations and fieldwork summary*. Dostopno na <http://ess.nsd.uib.no/index.jsp?year=2003&country=&module=fworksummary> (16. april 2008).

ESS DATA (2003c): *ESS Round 1, Fieldwork documents: Slovenia ESS1 – Main, Supplementary and Interviewer questionnaires*. Dostopno na <http://ess.nsd.uib.no/index.jsp?module=questionnaires&year=2003&countrySI=&download=%5CFieldwork+documents%5CSI%5C2003%5C01%23ESS1+-+Main%2C+Supplementary+and+Interviewer+questionnaires%5CLanguages%5CSlovenian%5CESS1QuestionnairesSI.rtf> (16. april 2008).

ESS DATA (2003d): *ESS Round 1, Survey documentation: ESS 1 Appendix 1, Population statistics and other documentation*. Dostopno na <http://ess.nsd.uib.no/index.jsp?module=documentation&year=2003&country=&download=%5CSurvey+documentation%5C2003%5C02%23ESS1+-+Appendix1%2C+Population+statistics+and+other+documentation%5CLanguages%5CEnglish%5CESS1AppendixA1.doc> (13. marec 2008).

ESS DATA (2003e): *ESS Round 1, Survey documentation: ESS 1 Appendix 3, Variables and questions, ed. 6.0*. Dostopno na [http://ess.nsd.uib.no/index.jsp?module=documentation&year=2003&country=&download=%5CSurvey+documentation%5C2003%5C04%23ESS1+-+Appendix3%2C+Variables+and+questions%2C+ed.+6.0%5CLanguages%5CEnglish%5CESS1AppendixA3\\_e06.doc](http://ess.nsd.uib.no/index.jsp?module=documentation&year=2003&country=&download=%5CSurvey+documentation%5C2003%5C04%23ESS1+-+Appendix3%2C+Variables+and+questions%2C+ed.+6.0%5CLanguages%5CEnglish%5CESS1AppendixA3_e06.doc) (16. april 2008).

ESS DATA (2005a): *ESS Round 2, Data download: ESS 2 – integrated file, edition 3.0*. Dostopno na <http://ess.nsd.uib.no/index.jsp?module=download&year=2005&country=&download=%5CDirect+Data+download%5C2005%5C01%23ESS2+-+integrated+file%2C+edition+3.0%5C.%5CESS2e03.spss.zip> (4. april 2008).

ESS DATA (2005b): *ESS Round 2, Deviations and fieldwork summary*. Dostopno na <http://ess.nsd.uib.no/index.jsp?year=2005&module=fworksummary&country=> (16. april 2008).

ESS DATA (2005c): *ESS Round 2, Survey documentation: ESS 2 Appendix 1, Population statistics and other documentation*, ed. 3.0. Dostopno na [http://ess.nsd.uib.no/index.jsp?module=documentation&year=2005&country=&download=%5CSurvey+documentation%5C2005%5C02%23ESS2+-+Appendix1%2C++Population+statistics+and+other+documentation%2C+ed.+3.0%5CLanguages%5CEnglish%5CESS2AppendixA1\\_e03.doc](http://ess.nsd.uib.no/index.jsp?module=documentation&year=2005&country=&download=%5CSurvey+documentation%5C2005%5C02%23ESS2+-+Appendix1%2C++Population+statistics+and+other+documentation%2C+ed.+3.0%5CLanguages%5CEnglish%5CESS2AppendixA1_e03.doc) (13. marec 2008).

ESS DATA (2005d): *ESS Round 2, Survey documentation: ESS 2 Appendix 3, Variables and questions*, ed. 3.0. Dostopno na [http://ess.nsd.uib.no/index.jsp?module=documentation&year=2005&country=&download=%5CSurvey+documentation%5C2005%5C04%23ESS2+-+Appendix3%2C+Variables+and+questions%2C+ed.+3.0%5CLanguages%5CEnglish%5CESS2AppendixA3\\_e03.doc](http://ess.nsd.uib.no/index.jsp?module=documentation&year=2005&country=&download=%5CSurvey+documentation%5C2005%5C04%23ESS2+-+Appendix3%2C+Variables+and+questions%2C+ed.+3.0%5CLanguages%5CEnglish%5CESS2AppendixA3_e03.doc) (16. april 2008).

ESS DATA (2007a): *ESS Round 3, Data download: ESS 3 – integrated file, edition 3.1*. Dostopno na [http://ess.nsd.uib.no/index.jsp?module=download&year=2007&country=&download=%5CDirect+Data+download%5C2007%5C02%23ESS3+-+Latvia+and+Romania+%28no+design+weights%29%5C.%5CESS3LVRO.spss.zip](http://ess.nsd.uib.no/index.jsp?module=download&year=2007&country=&download=%5CDirect+Data+download%5C2007%5C01%23ESS3+-+integrated+file%2C+edition+3.1+%28Latvia+and++Romania+not+included%29%5C.%5CESS3e03_1.spss.zip) in <http://ess.nsd.uib.no/index.jsp?module=download&year=2007&country=&download=%5CDirect+Data+download%5C2007%5C02%23ESS3+-+Latvia+and+Romania+%28no+design+weights%29%5C.%5CESS3LVRO.spss.zip> (4. april 2008).

ESS DATA (2007b): *ESS Round 3, Deviations and fieldwork summary*. Dostopno na <http://ess.nsd.uib.no/index.jsp?year=2007&module=fworksummary&country=> (16. april 2008).

ESS DATA (2007c): *ESS Round 3, Survey documentation: ESS 3 Appendix 1, Population statistics*, ed. 3.1. Dostopno na [http://ess.nsd.uib.no/index.jsp?module=documentation&year=2007&country=&download=%5CSurvey+documentation%5C2007%5C02%23ESS3+-+Appendix1%2C++Population+statistics%2C+ed.+3.1%5CLanguages%5CEnglish%5CESS3AppendixA1\\_e03\\_1.doc](http://ess.nsd.uib.no/index.jsp?module=documentation&year=2007&country=&download=%5CSurvey+documentation%5C2007%5C02%23ESS3+-+Appendix1%2C++Population+statistics%2C+ed.+3.1%5CLanguages%5CEnglish%5CESS3AppendixA1_e03_1.doc) (14. april 2008).

ESS DATA (2007d): *ESS Round 3, Survey documentation: ESS 3 Appendix 3, Variables and questions*, ed. 3.1. Dostopno na [http://ess.nsd.uib.no/index.jsp?module=documentation&year=2007&country=&download=%5CSurvey+documentation%5C2007%5C04%23ESS3+-+Appendix3%2C+Variables+and+questions%2C+ed.+3.1%5CLanguages%5CEnglish%5CESS3AppendixA3\\_e03\\_1.doc](http://ess.nsd.uib.no/index.jsp?module=documentation&year=2007&country=&download=%5CSurvey+documentation%5C2007%5C04%23ESS3+-+Appendix3%2C+Variables+and+questions%2C+ed.+3.1%5CLanguages%5CEnglish%5CESS3AppendixA3_e03_1.doc) (16. april 2008).

EU (2008 zadnja sprememba): *Basic principles of the NUTS*. Dostopno na [http://ec.europa.eu/eurostat/ramon/nuts/basicnuts\\_regions\\_en.html](http://ec.europa.eu/eurostat/ramon/nuts/basicnuts_regions_en.html) (14. marec 2008).

European Communities (2008): *Hierarchical list of the Nomenclature of territorial units for statistics – NUTS and the Statistical regions of Europe*. Dostopno na [http://ec.europa.eu/eurostat/ramon/nuts/codelist\\_en.cfm](http://ec.europa.eu/eurostat/ramon/nuts/codelist_en.cfm) (14. marec 2008).

European social survey (2007a): *Home*. Dostopno na <http://www.europeansocialsurvey.org/> (16. april 2008).

- European social survey (2007b): *Sampling designs*. Dostopno na [http://www.europeansocialsurvey.org/index.php?option=com\\_content&task=view&id=80&Itemid=125](http://www.europeansocialsurvey.org/index.php?option=com_content&task=view&id=80&Itemid=125) (16. april 2008).
- European social survey (2007c): *ESS Round 1 2002/2003 Technical Report Edition 2 (June 2004)*. Dostopno na [http://www.europeansocialsurvey.org/index.php?option=com\\_content&task=view&id=101&Itemid=139](http://www.europeansocialsurvey.org/index.php?option=com_content&task=view&id=101&Itemid=139) (16. april 2008).
- European social survey (2007d): *ESS Round 2 2004/2005 Technical Report Edition 1 (June 2007)*. Dostopno na [http://www.europeansocialsurvey.org/index.php?option=com\\_content&task=view&id=101&Itemid=139](http://www.europeansocialsurvey.org/index.php?option=com_content&task=view&id=101&Itemid=139) (16. april 2008).
- Eurostat (2007a): *Developments and organisation of the EU Labour force survey*. Dostopno na [http://circa.europa.eu/irc/dsis/employment/info/data/eu\\_lfs/LFS\\_MAIN/LFS\\_DEVELOPMENT.htm](http://circa.europa.eu/irc/dsis/employment/info/data/eu_lfs/LFS_MAIN/LFS_DEVELOPMENT.htm) (16. april 2008).
- Eurostat (2007b): *Levels of education and training in the EU Labour Force Survey*. Dostopno na [http://circa.europa.eu/irc/dsis/employment/info/data/eu\\_lfs/LFS\\_MAIN/Related\\_documents/ISCED\\_EN.htm](http://circa.europa.eu/irc/dsis/employment/info/data/eu_lfs/LFS_MAIN/Related_documents/ISCED_EN.htm) (14. marec 2008).
- Eurostat (2007c): LFS Series – Quarterly survey results. Dostopno na [http://circa.europa.eu/irc/dsis/employment/info/data/eu\\_lfs/LFS\\_MAIN/Related\\_documents/lfs\\_series\\_sm.doc](http://circa.europa.eu/irc/dsis/employment/info/data/eu_lfs/LFS_MAIN/Related_documents/lfs_series_sm.doc) (14. marec 2008).
- Eurostat (2007d): *Statistical References: Brief information on Eurostat Products and services*. Luxembourg: Office for Official Publications of the European Communities. Dostopno na [http://epp.eurostat.ec.europa.eu/cache/ITY\\_OFFPUB/KS-CL-07-001/EN/KS-CL-07-001-EN.PDF](http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-CL-07-001/EN/KS-CL-07-001-EN.PDF) (17. april 2008).
- Eurostat (2008a): *LFS agregirani podatki za leta 2002, 2004, 2006*. Interno gradivo.
- Eurostat (2008b): *Regional codes: NUTS II*. Dostopno na [http://circa.europa.eu/irc/dsis/employment/info/data/eu\\_lfs/LFS\\_MAIN/Related\\_documents/Info\\_users\\_Regions\\_EN.xls](http://circa.europa.eu/irc/dsis/employment/info/data/eu_lfs/LFS_MAIN/Related_documents/Info_users_Regions_EN.xls) (14. marec 2008).
- Eurostat (2008c): *What is Eurostat*. Dostopno na [http://epp.eurostat.ec.europa.eu/portal/page?\\_pageid=1393,47358531,1393\\_47358596&\\_dad=portal&\\_schema=PORTAL](http://epp.eurostat.ec.europa.eu/portal/page?_pageid=1393,47358531,1393_47358596&_dad=portal&_schema=PORTAL) (17. april 2008).
- Fowler, Floyd J. (2002): *Survey research methods*. Neybury Park, London, New Delhi: SAGE publications.
- Ganninger, Matthias (2007): *Weighting ESS cumulative data set*. Dostopno na <http://ess.nsd.uib.no/index.jsp?module=main&year=-3&country=&download=%5CCumulative+Documentation%5C02%23ESS1-2+-+Weighting+ESS+cumulative+data%5C.%5CWeighting+ESS+cumulative+data+pdf.pdf> (13. avgust 2008).
- Groves, Robert M. (2006): Nonresponde rates and nonresponse bias in household surveys. *Public Opinion Quarterly* 70 (5), 646–657.

- Groves, Robert M., Floyd J. Fowler, Mick P. Couper, James M. Lepkowski, Eleanor Singer in Roger Tourangeau (2004): *Survey methodology*. Hoboken, New Jersey: John Wiley & Sons, Inc.
- Häder, Sabine in Peter Lynn (2007): How representative can a multi-nation survey be? V Roger Jowell, Caroline Roberts, Rory Fitzgerald in Gillian Eva (ur.): *Measuring attitudes cross-nationally: lessons from the European Social Survey*, 33–52. Los Angeles: Sage Publications.
- Jowell, Roger in Brina Malnar (2004): *Evropska družboslovna raziskava 2002/2003*. Ljubljana: Fakulteta za družbene vede, Arhiv družboslovnih podatkov. <http://www.adp.fdv.uni-lj.si/opisi/ess03.xml> (8. maj 2008).
- Jowell, Roger in Brina Malnar (2006): *Evropska družboslovna raziskava 2004 = European social survey 2004*. Ljubljana: Fakulteta za družbene vede, Arhiv družboslovnih podatkov. <http://www.adp.fdv.uni-lj.si/opisi/ess04.xml> (8. maj 2008).
- Jowell, Roger, Max Kaase, Rory Fitzgerald in Gillian Eva (2007): The European Social Survey as a measurement model. V Roger Jowell, Caroline Roberts, Rory Fitzgerald in Gillian Eva (ur.): *Measuring attitudes cross-nationally: lessons from the European Social Survey*, 1–31. Los Angeles: Sage Publications.
- Jowell, Roger, Peter Mohler, Ineke Stoop, Willem Saris, Jaak Billiet in Bjorn Henrichsen (2006): *European Social Survey: Round 2, End of grant report*. Dostopno na [http://www.europeansocialsurvey.org/index.php?option=com\\_docman&task= doc\\_view&gid=76](http://www.europeansocialsurvey.org/index.php?option=com_docman&task= doc_view&gid=76) (21. april 2008).
- Kalton, Graham (1983): *Introduction to survey sampling*. Neybury Park, London, New Delhi: SAGE publications.
- Kalton, Graham in Vasja Vehovar (2001): *Vzorčenje v anketah*. Ljubljana: Fakulteta za družbene vede.
- Kolsrud, Kirstine, Knut Kalgraff Skjåk in Bjørn Henrichsen (2007): Free and immediate access to data. V Roger Jowell, Caroline Roberts, Rory Fitzgerald in Gillian Eva (ur.): *Measuring attitudes cross-nationally: lessons from the European Social Survey*, 139–156. Los Angeles: Sage Publications.
- Kveder, Andrej (2001): *Neodgovori na anketna vprašanja: magistrsko delo*. Ljubljana: [A. Kveder].
- LFS (2003): *The European Union Labour Force Survey – Methods and definitions 2001*. Luxembourg: Office for Official Publications of the European Communities. Dostopno na [http://epp.eurostat.ec.europa.eu/cache/ITY\\_OFFPUB/KS-BF-03-002/EN/KS-BF-03-002-EN.PDF](http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-BF-03-002/EN/KS-BF-03-002-EN.PDF) (14. marec 2008).
- Malnar, Brina (2004): *Evropska družboslovna raziskava 2002: Zaključno raziskovalno poročilo* (CRP Konkurenčnost Slovenije 2001–2006). Ljubljana: Fakulteta za družbene vede, IDV, Center za raziskovanje javnega mnenja in množičnih komunikacij. Dostopno na [http://cjm.si/sites/cjm.si/files/File/e-dokumenti/aESS2002\\_ELABORAT\\_mega.pdf](http://cjm.si/sites/cjm.si/files/File/e-dokumenti/aESS2002_ELABORAT_mega.pdf) (8. maj 2008).

- Mohler, Peter (2007): What is being learned from the ESS. V Roger Jowell, Caroline Roberts, Rory Fitzgerald in Gillian Eva (ur.): *Measuring attitudes cross-nationally: lessons from the European Social Survey*, 157–168. Los Angeles: Sage Publications.
- Norwegian Social Science Data Services (2008a): *European Social Survey Education net (ESS EduNet)*. Dostopno na <http://essedunet.nsd.uib.no/cms/edunet/about.html> (15. april 2008).
- Norwegian Social Science Data Services (2008b): *European Social Survey Education net (ESS EduNet): Weighting the ESS data*. Dostopno na <http://essedunet.nsd.uib.no/cms/userguide/weight/> (16. april 2008).
- Rydland, Lars Tore; Sveinung Arnesen in Åse Gilje Østensen (2007): *ESS Report on Contextual Data*. Dostopno na [http://ess.nsd.uib.no/files/ESS\\_context\\_sources.pdf](http://ess.nsd.uib.no/files/ESS_context_sources.pdf) (4. april 2008).
- Statistični urad (2008): *Klasifikacija statističnih teritorialnih enot v Evropski uniji – NUTS*. Dostopno na <http://www.stat.si/doc/reg/NUTS%20pojasnilo.doc> (16. april 2008).
- Stoop, Ineke (2007): Survey data, context and event data. V Mike Trotman, Tim Burrell, Laurance Gerrard, Ken Anderton, Garj Basi, Mick Couper, Keigh Morris, David Birks, AJ Johnson, Reg Baker, Malcolm Rigg, Steve Taylor in Andrew Westlake (ur.): *ASC 2007. The challenges of a changing world. Association for Survey Computing*, 17–28. Dostopno na <http://www2.asc.org.uk/ASC/Sep2007/Conference/files/papers/ASC19-event%20paper-Stoop-reviewd-July.pdf> (21. april 2008).
- Vehovar, Vasja in Tina Zupanič (2007). *Weighting in the ESS – Round 2*. Fakulteta za družbene vede. Interno gradivo.
- Vehovar, Vasja (2006). *Weighting in ESS – Round 1*. Fakulteta za družbene vede. Interno gradivo.

# Priloge

## **Priloga A: Kategorizacija NUTS za vse tri obravnavane države**

**Tabela A1: NUTS-kategorizacija za BELGIJO**

NUTS-koda	nivo NUTS 1	nivo NUTS 2	nivo NUTS 3
<b>BE</b>			
BE1	RÉGION DE BRUXELLES-CAPITALE		
BE10		Région de Bruxelles-Capitale	
BE100			Arr. de Bruxelles-Capitale
BE2	VLAAMS GEWEST		
BE21		Prov. Antwerpen	
BE211			Arr. Antwerpen
BE212			Arr. Mechelen
BE213			Arr. Turnhout
BE22		Prov. Limburg (B)	
BE221			Arr. Hasselt
BE222			Arr. Maaseik
BE223			Arr. Tongeren
BE23		Prov. Oost-Vlaanderen	
BE231			Arr. Aalst
BE232			Arr. Dendermonde
BE233			Arr. Eeklo
BE234			Arr. Gent
BE235			Arr. Oudenaarde
BE236			Arr. Sint-Niklaas
BE24		Prov. Vlaams-Brabant	
BE241			Arr. Halle-Vilvoorde
BE242			Arr. Leuven
BE25		Prov. West-Vlaanderen	
BE251			Arr. Brugge
BE252			Arr. Diksmuide
BE253			Arr. Ieper
BE254			Arr. Kortrijk
BE255			Arr. Oostende
BE256			Arr. Roeselare
BE257			Arr. Tielt
BE258			Arr. Veurne
BE3	RÉGION WALLONNE		
BE31		Prov. Brabant Wallon	
BE310			Arr. Nivelles
BE32		Prov. Hainaut	
BE321			Arr. Ath
BE322			Arr. Charleroi
BE323			Arr. Mons
BE324			Arr. Mouscron
BE325			Arr. Soignies
BE326			Arr. Thuin
BE327			Arr. Tournai
BE33		Prov. Liège	
BE331			Arr. Huy
BE332			Arr. Liège
BE334			Arr. Waremme

BE335			Arr. Verviers - communes francophones
BE336			Bezirk Verviers - Deutschsprachige G.
BE34		Prov. Luxembourg (B)	
BE341			Arr. Arlon
BE342			Arr. Bastogne
BE343			Arr. Marche-en-Famenne
BE344			Arr. Neufchâteau
BE345			Arr. Virton
BE35		Prov. Namur	
BE351			Arr. Dinant
BE352			Arr. Namur
BE353			Arr. Philippeville
BEZ	EXTRA-REGIO		
BEZZ		Extra-Regio	
BEZZZ			Extra-Regio

**Tabela A2: NUTS-kategorizacija za SLOVENIJO**

NUTS-koda	nivo NUTS 1	nivo NUTS 2	nivo NUTS 3
<b>SI</b>			
SI0	SLOVENIJA		
SI01		Vzhodna Slovenija	
SI011			Pomurska
SI012			Podravska
SI013			Koroska
SI014			Savinjska
SI015			Zasavska
SI016			Spodnjeposavska
SI017			Jugovzhodna Slovenija
SI018			Notranjsko-kraska
SI02		Zahodna Slovenija	
SI021			Osrednjeslovenska
SI022			Gorenjska
SI023			Goriska
SI024			Obalno-kraska
SIZ	EXTRA-REGIO		
SIZZ		Extra-Regio	
SIZZZ			Extra-Regio

**Tabela A3: NUTS-kategorizacija za ŠVEDSKO**

NUTS-koda	nivo NUTS 1	nivo NUTS 2	nivo NUTS 3
<b>SE</b>			
SE1	ÖSTRA SVERIGE		
SE11		Stockholm	
SE110			Stockholms län
SE12		Östra Mellansverige	
SE121			Uppsala län
SE122			Södermanlands län
SE123			Östergötlands län
SE124			Örebro län
SE125			Västmanlands län
SE2	SÖDRA SVERIGE		
SE21		Småland med öarna	
SE211			Jönköpings län

SE212			Kronobergs län
SE213			Kalmar län
SE214			Gotlands län
SE22		Sydsverige	
SE221			Blekinge län
SE224			Skåne län
SE23		Västsverige	
SE231			Hallands län
SE232			Västra Götalands län
SE3	NORRA SVERIGE		
SE31		Norra Mellansverige	
SE311			Värmlands län
SE312			Dalarnas län
SE313			Gävleborgs län
SE32		Mellersta Norrland	
SE321			Västernorrlands län
SE322			Jämtlands län
SE33		Övre Norrland	
SE331			Västerbottens län
SE332			Norrbottens län
SEZ	EXTRA-REGIO		
SEZZ		Extra-Regio	
SEZZZ			Extra-Regio

























## **Priloga C: Primer sintakse za raking, Belgija R3**

```
FREQUENCIES
  VARIABLES=gndr age edulvl regionbe
  /ORDER= ANALYSIS .

RECODE
  gndr
  (MISSING=0) (ELSE=Copy) INTO gndr_r .
  VARIABLE LABELS gndr_r 'gender with missings'.
  VALUE LABELS gndr_r 0 'Missing' 1 'Male' 2 'Female'.
  EXECUTE .

RECODE
  age
  (MISSING=0) (Lowest thru 34.99=1) (35 thru 54.99=2) (55 thru Highest=3)
  INTO age_3gr .
  VARIABLE LABELS age_3gr 'age with missings'.
  VALUE LABELS age_3gr 0 'Missing' 1 '15-34' 2 '35-54' 3 '55+'.
  EXECUTE .

RECODE
  edulvl
  (0 thru 2=1) (3 thru 4=2) (5 thru 6=3) (ELSE=0) INTO edu3grB .
  VARIABLE LABELS edu3grB 'education 3 groups'.
  VALUE LABELS edu3grB 0 'Missing' 1 'Low' 2 'Medium' 3 'High'.
  EXECUTE .

CROSSTABS
  /TABLES=age_3gr BY edu3grB BY gndr_r
  /FORMAT= AVALUE TABLES
  /CELLS= COUNT
  /COUNT ASIS .

COMPUTE w_psN = dweight .
EXECUTE .
COMPUTE w_psN = w_psN* 8749360.58/1798 .
EXECUTE .

**ponavlja**

WEIGHT
  BY w_psN .
CROSSTABS
  /TABLES=age_3gr BY edu3grB BY gndr_r
  /FORMAT= AVALUE TABLES
  /CELLS= COUNT
  /COUNT ASIS .
FREQUENCIES
  VARIABLES=regionbe
  /ORDER= ANALYSIS .
WEIGHT
  OFF.

*gae1.
IF (gndr_r = 1 & age_3gr = 1 & edu3grB = 1) w_psN = w_psN*466418.875812013/389293.018020022.
IF (gndr_r = 1 & age_3gr = 1 & edu3grB = 2) w_psN = w_psN*563566.49980812/564474.876129033.
IF (gndr_r = 1 & age_3gr = 1 & edu3grB = 3) w_psN = w_psN*303931.972016129/321166.739866518.
IF (gndr_r = 1 & age_3gr = 2 & edu3grB = 1) w_psN = w_psN*524611.232753059/287103.600789766.
IF (gndr_r = 1 & age_3gr = 2 & edu3grB = 2) w_psN = w_psN*564319.670681312/618002.666106785.
IF (gndr_r = 1 & age_3gr = 2 & edu3grB = 3) w_psN = w_psN*469373.096838153/530411.737052281.
IF (gndr_r = 1 & age_3gr = 2 & edu3grB = 0) w_psN = w_psN*4866.16272525028/4866.16272525028.
IF (gndr_r = 1 & age_3gr = 3 & edu3grB = 1) w_psN = w_psN*767806.411263904/705593.595161289.
IF (gndr_r = 1 & age_3gr = 3 & edu3grB = 2) w_psN = w_psN*325888.813406563/413623.831646274.
IF (gndr_r = 1 & age_3gr = 3 & edu3grB = 3) w_psN = w_psN*255564.427733593/253040.461713014.
```

```

IF (gndr_r = 2 & age_3gr = 1 & edu3grB = 1) w_psN = w_psN*389035.693576249/437954.645272525.
IF (gndr_r = 2 & age_3gr = 1 & edu3grB = 2) w_psN = w_psN*515288.250780406/540144.062502781.
IF (gndr_r = 2 & age_3gr = 1 & edu3grB = 3) w_psN = w_psN*403802.605420875/364962.204393771.
IF (gndr_r = 2 & age_3gr = 2 & edu3grB = 1) w_psN = w_psN*491722.809602336/394159.180745273.
IF (gndr_r = 2 & age_3gr = 2 & edu3grB = 2) w_psN = w_psN*554526.195583982/656931.967908787.
IF (gndr_r = 2 & age_3gr = 2 & edu3grB = 3) w_psN = w_psN*485148.732964405/705593.595161289.
IF (gndr_r = 2 & age_3gr = 3 & edu3grB = 1) w_psN = w_psN*1122471.68605812/924570.91779755.
IF (gndr_r = 2 & age_3gr = 3 & edu3grB = 2) w_psN = w_psN*325506.153850111/379560.692569522.
IF (gndr_r = 2 & age_3gr = 3 & edu3grB = 3) w_psN = w_psN*215511.289125417/257906.624438264.
EXECUTE .

```

\*region1.

```

IF (regionbe = 1) w_psN = w_psN*5102865.9875/5520973.35999658.
IF (regionbe = 2) w_psN = w_psN*835237.0075/459309.941338693.
IF (regionbe = 3) w_psN = w_psN*2811257.585/2769077.27866473.
EXECUTE .

```

\*gae2.

```

IF (gndr_r = 1 & age_3gr = 1 & edu3grB = 1) w_psN = w_psN*466418.875812013/474574.478507708.
IF (gndr_r = 1 & age_3gr = 1 & edu3grB = 2) w_psN = w_psN*563566.49980812/574569.392512145.
IF (gndr_r = 1 & age_3gr = 1 & edu3grB = 3) w_psN = w_psN*303931.972016129/309462.758711517.
IF (gndr_r = 1 & age_3gr = 2 & edu3grB = 1) w_psN = w_psN*524611.232753059/506582.878508108.
IF (gndr_r = 1 & age_3gr = 2 & edu3grB = 2) w_psN = w_psN*564319.670681312/550814.890075786.
IF (gndr_r = 1 & age_3gr = 2 & edu3grB = 3) w_psN = w_psN*469373.096838153/492176.487691905.
IF (gndr_r = 1 & age_3gr = 2 & edu3grB = 0) w_psN = w_psN*4866.16272525028/4940.28714063218.
IF (gndr_r = 1 & age_3gr = 3 & edu3grB = 1) w_psN = w_psN*767806.411263904/739841.579481254.
IF (gndr_r = 1 & age_3gr = 3 & edu3grB = 2) w_psN = w_psN*325888.813406563/313007.384593242.
IF (gndr_r = 1 & age_3gr = 3 & edu3grB = 3) w_psN = w_psN*255564.427733593/268390.189509242.
IF (gndr_r = 2 & age_3gr = 1 & edu3grB = 1) w_psN = w_psN*389035.693576249/405830.515772183.
IF (gndr_r = 2 & age_3gr = 1 & edu3grB = 2) w_psN = w_psN*515288.250780406/509760.21885809.
IF (gndr_r = 2 & age_3gr = 1 & edu3grB = 3) w_psN = w_psN*403802.605420875/409048.118797814.
IF (gndr_r = 2 & age_3gr = 2 & edu3grB = 1) w_psN = w_psN*491722.809602336/500400.93399987.
IF (gndr_r = 2 & age_3gr = 2 & edu3grB = 2) w_psN = w_psN*554526.195583982/547710.280818117.
IF (gndr_r = 2 & age_3gr = 2 & edu3grB = 3) w_psN = w_psN*485148.732964405/501301.992471018.
IF (gndr_r = 2 & age_3gr = 3 & edu3grB = 1) w_psN = w_psN*1122471.68605812/1089227.83652749.
IF (gndr_r = 2 & age_3gr = 3 & edu3grB = 2) w_psN = w_psN*325506.153850111/333114.641609457.
IF (gndr_r = 2 & age_3gr = 3 & edu3grB = 3) w_psN = w_psN*215511.289125417/218605.714414418.
EXECUTE .

```

\*region2.

```

IF (regionbe = 1) w_psN = w_psN*5102865.9875/5117995.0966497.
IF (regionbe = 2) w_psN = w_psN*835237.0075/821391.835423468.
IF (regionbe = 3) w_psN = w_psN*2811257.585/2809973.64792683.
EXECUTE .

```

\*gae3.

```

IF (gndr_r = 1 & age_3gr = 1 & edu3grB = 1) w_psN = w_psN*466418.875812013/466735.38208552.
IF (gndr_r = 1 & age_3gr = 1 & edu3grB = 2) w_psN = w_psN*563566.49980812/563990.559171511.
IF (gndr_r = 1 & age_3gr = 1 & edu3grB = 3) w_psN = w_psN*303931.972016129/304145.983068553.
IF (gndr_r = 1 & age_3gr = 2 & edu3grB = 1) w_psN = w_psN*524611.232753059/523934.593715613.
IF (gndr_r = 1 & age_3gr = 2 & edu3grB = 2) w_psN = w_psN*564319.670681312/563828.085495328.
IF (gndr_r = 1 & age_3gr = 2 & edu3grB = 3) w_psN = w_psN*469373.096838153/470202.122931324.
IF (gndr_r = 1 & age_3gr = 2 & edu3grB = 0) w_psN = w_psN*4866.16272525028/4868.38617910069.
IF (gndr_r = 1 & age_3gr = 3 & edu3grB = 1) w_psN = w_psN*767806.411263904/766751.466840353.
IF (gndr_r = 1 & age_3gr = 3 & edu3grB = 2) w_psN = w_psN*325888.813406563/325401.23019737.
IF (gndr_r = 1 & age_3gr = 3 & edu3grB = 3) w_psN = w_psN*255564.427733593/256031.368003014.
IF (gndr_r = 2 & age_3gr = 1 & edu3grB = 1) w_psN = w_psN*389035.693576249/389639.928397217.
IF (gndr_r = 2 & age_3gr = 1 & edu3grB = 2) w_psN = w_psN*515288.250780406/515091.65097604.
IF (gndr_r = 2 & age_3gr = 1 & edu3grB = 3) w_psN = w_psN*403802.605420875/404008.346832972.
IF (gndr_r = 2 & age_3gr = 2 & edu3grB = 1) w_psN = w_psN*491722.809602336/492046.593589075.
IF (gndr_r = 2 & age_3gr = 2 & edu3grB = 2) w_psN = w_psN*554526.195583982/554284.555236414.
IF (gndr_r = 2 & age_3gr = 2 & edu3grB = 3) w_psN = w_psN*485148.732964405/485747.249999417.
IF (gndr_r = 2 & age_3gr = 3 & edu3grB = 1) w_psN = w_psN*1122471.68605812/1121233.33869634.
IF (gndr_r = 2 & age_3gr = 3 & edu3grB = 2) w_psN = w_psN*325506.153850111/325792.780726275.

```

IF (gndr\_r = 2 & age\_3gr = 3 & edu3grB = 3) w\_psN = w\_psN\*215511.289125417/215626.957858574.  
EXECUTE .

\*region3.

IF (regionbe = 1) w\_psN = w\_psN\*5102865.9875/5103430.75012238.  
IF (regionbe = 2) w\_psN = w\_psN\*835237.0075/834715.586678161.  
IF (regionbe = 3) w\_psN = w\_psN\*2811257.585/2811214.24319944.  
EXECUTE .

\*gae4.

IF (gndr\_r = 1 & age\_3gr = 1 & edu3grB = 1) w\_psN = w\_psN\*466418.875812013/466430.841566039.  
IF (gndr\_r = 1 & age\_3gr = 1 & edu3grB = 2) w\_psN = w\_psN\*563566.49980812/563582.527543889.  
IF (gndr\_r = 1 & age\_3gr = 1 & edu3grB = 3) w\_psN = w\_psN\*303931.972016129/303940.060955292.  
IF (gndr\_r = 1 & age\_3gr = 2 & edu3grB = 1) w\_psN = w\_psN\*524611.232753059/524585.786334314.  
IF (gndr\_r = 1 & age\_3gr = 2 & edu3grB = 2) w\_psN = w\_psN\*564319.670681312/564301.257736338.  
IF (gndr\_r = 1 & age\_3gr = 2 & edu3grB = 3) w\_psN = w\_psN\*469373.096838153/469404.316938783.  
IF (gndr\_r = 1 & age\_3gr = 2 & edu3grB = 0) w\_psN = w\_psN\*4866.16272525028/4866.23774914959.  
IF (gndr\_r = 1 & age\_3gr = 3 & edu3grB = 1) w\_psN = w\_psN\*767806.411263904/767766.710315045.  
IF (gndr\_r = 1 & age\_3gr = 3 & edu3grB = 2) w\_psN = w\_psN\*325888.813406563/325870.47208205.  
IF (gndr\_r = 1 & age\_3gr = 3 & edu3grB = 3) w\_psN = w\_psN\*255564.427733593/255582.032555306.  
IF (gndr\_r = 2 & age\_3gr = 1 & edu3grB = 1) w\_psN = w\_psN\*389035.693576249/389058.30902541.  
IF (gndr\_r = 2 & age\_3gr = 1 & edu3grB = 2) w\_psN = w\_psN\*515288.250780406/515280.805629746.  
IF (gndr\_r = 2 & age\_3gr = 1 & edu3grB = 3) w\_psN = w\_psN\*403802.605420875/403810.362122872.  
IF (gndr\_r = 2 & age\_3gr = 2 & edu3grB = 1) w\_psN = w\_psN\*491722.809602336/491734.879508005.  
IF (gndr\_r = 2 & age\_3gr = 2 & edu3grB = 2) w\_psN = w\_psN\*554526.195583982/554517.09133988.  
IF (gndr\_r = 2 & age\_3gr = 2 & edu3grB = 3) w\_psN = w\_psN\*485148.732964405/485171.261689217.  
IF (gndr\_r = 2 & age\_3gr = 3 & edu3grB = 1) w\_psN = w\_psN\*1122471.68605812/1122425.1064431.  
IF (gndr\_r = 2 & age\_3gr = 3 & edu3grB = 2) w\_psN = w\_psN\*325506.153850111/325516.936543428.  
IF (gndr\_r = 2 & age\_3gr = 3 & edu3grB = 3) w\_psN = w\_psN\*215511.289125417/215515.583922145.  
EXECUTE .

\*region4.

IF (regionbe = 1) w\_psN = w\_psN\*5102865.9875/5102887.1937321.  
IF (regionbe = 2) w\_psN = w\_psN\*835237.0075/835217.37487266.  
IF (regionbe = 3) w\_psN = w\_psN\*2811257.585/2811256.01139521.  
EXECUTE .

\*gae5.

IF (gndr\_r = 1 & age\_3gr = 1 & edu3grB = 1) w\_psN = w\_psN\*466418.875812013/466419.326677147.  
IF (gndr\_r = 1 & age\_3gr = 1 & edu3grB = 2) w\_psN = w\_psN\*563566.49980812/563567.103719144.  
IF (gndr\_r = 1 & age\_3gr = 1 & edu3grB = 3) w\_psN = w\_psN\*303931.972016129/303932.276789864.  
IF (gndr\_r = 1 & age\_3gr = 2 & edu3grB = 1) w\_psN = w\_psN\*524611.232753059/524610.275287226.  
IF (gndr\_r = 1 & age\_3gr = 2 & edu3grB = 2) w\_psN = w\_psN\*564319.670681312/564318.978546512.  
IF (gndr\_r = 1 & age\_3gr = 2 & edu3grB = 3) w\_psN = w\_psN\*469373.096838153/469374.272490726.  
IF (gndr\_r = 1 & age\_3gr = 2 & edu3grB = 0) w\_psN = w\_psN\*4866.16272525028/4866.1654490922.  
IF (gndr\_r = 1 & age\_3gr = 3 & edu3grB = 1) w\_psN = w\_psN\*767806.411263904/767804.91718869.  
IF (gndr\_r = 1 & age\_3gr = 3 & edu3grB = 2) w\_psN = w\_psN\*325888.813406563/325888.12328914.  
IF (gndr\_r = 1 & age\_3gr = 3 & edu3grB = 3) w\_psN = w\_psN\*255564.427733593/255565.090922565.  
IF (gndr\_r = 2 & age\_3gr = 1 & edu3grB = 1) w\_psN = w\_psN\*389035.693576249/389036.543558261.  
IF (gndr\_r = 2 & age\_3gr = 1 & edu3grB = 2) w\_psN = w\_psN\*515288.250780406/515287.969830237.  
IF (gndr\_r = 2 & age\_3gr = 1 & edu3grB = 3) w\_psN = w\_psN\*403802.605420875/403802.897390006.  
IF (gndr\_r = 2 & age\_3gr = 2 & edu3grB = 1) w\_psN = w\_psN\*491722.809602336/491723.26245012.  
IF (gndr\_r = 2 & age\_3gr = 2 & edu3grB = 2) w\_psN = w\_psN\*554526.195583982/554525.85258998.  
IF (gndr\_r = 2 & age\_3gr = 2 & edu3grB = 3) w\_psN = w\_psN\*485148.732964405/485149.581022269.  
IF (gndr\_r = 2 & age\_3gr = 3 & edu3grB = 1) w\_psN = w\_psN\*1122471.68605812/1122469.93315855.  
IF (gndr\_r = 2 & age\_3gr = 3 & edu3grB = 2) w\_psN = w\_psN\*325506.153850111/325506.559598164.  
IF (gndr\_r = 2 & age\_3gr = 3 & edu3grB = 3) w\_psN = w\_psN\*215511.289125417/215511.450042335.  
EXECUTE .

\*region5.

IF (regionbe = 1) w\_psN = w\_psN\*5102865.9875/5102866.78505546.  
IF (regionbe = 2) w\_psN = w\_psN\*835237.0075/835236.268523017.  
IF (regionbe = 3) w\_psN = w\_psN\*2811257.585/2811257.52642152.  
EXECUTE .

\*gae6.

```

IF (gndr_r = 1 & age_3gr = 1 & edu3grB = 1) w_psN = w_psN*466418.875812013/466418.892786089.
IF (gndr_r = 1 & age_3gr = 1 & edu3grB = 2) w_psN = w_psN*563566.49980812/563566.522543998.
IF (gndr_r = 1 & age_3gr = 1 & edu3grB = 3) w_psN = w_psN*303931.972016129/303931.983490034.
IF (gndr_r = 1 & age_3gr = 2 & edu3grB = 1) w_psN = w_psN*524611.232753059/524611.196721351.
IF (gndr_r = 1 & age_3gr = 2 & edu3grB = 2) w_psN = w_psN*564319.670681312/564319.644642057.
IF (gndr_r = 1 & age_3gr = 2 & edu3grB = 3) w_psN = w_psN*469373.096838153/469373.141091892.
IF (gndr_r = 1 & age_3gr = 2 & edu3grB = 0) w_psN = w_psN*4866.16272525028/4866.16282664705.
IF (gndr_r = 1 & age_3gr = 3 & edu3grB = 1) w_psN = w_psN*767806.411263904/767806.355035516.
IF (gndr_r = 1 & age_3gr = 3 & edu3grB = 2) w_psN = w_psN*325888.813406563/325888.787435969.
IF (gndr_r = 1 & age_3gr = 3 & edu3grB = 3) w_psN = w_psN*255564.427733593/255564.452700016.
IF (gndr_r = 2 & age_3gr = 1 & edu3grB = 1) w_psN = w_psN*389035.693576249/389035.725552691.
IF (gndr_r = 2 & age_3gr = 1 & edu3grB = 2) w_psN = w_psN*515288.250780406/515288.240198193.
IF (gndr_r = 2 & age_3gr = 1 & edu3grB = 3) w_psN = w_psN*403802.605420875/403802.616409437.
IF (gndr_r = 2 & age_3gr = 2 & edu3grB = 1) w_psN = w_psN*491722.809602336/491722.826629394.
IF (gndr_r = 2 & age_3gr = 2 & edu3grB = 2) w_psN = w_psN*554526.195583982/554526.182671163.
IF (gndr_r = 2 & age_3gr = 2 & edu3grB = 3) w_psN = w_psN*485148.732964405/485148.764883224.
IF (gndr_r = 2 & age_3gr = 3 & edu3grB = 1) w_psN = w_psN*1122471.68605812/1122471.62008922.
IF (gndr_r = 2 & age_3gr = 3 & edu3grB = 2) w_psN = w_psN*325506.153850111/325506.169119699.
IF (gndr_r = 2 & age_3gr = 3 & edu3grB = 3) w_psN = w_psN*215511.289125417/215511.295173408.
EXECUTE .

```

\*region6.

```

IF (regionbe = 1) w_psN = w_psN*5102865.9875/5102866.01750987.
IF (regionbe = 2) w_psN = w_psN*835237.0075/835236.979687587.
IF (regionbe = 3) w_psN = w_psN*2811257.585/2811257.58280257.
EXECUTE .

```

\*gae7.

```

IF (gndr_r = 1 & age_3gr = 1 & edu3grB = 1) w_psN = w_psN*466418.875812013/466418.876450891.
IF (gndr_r = 1 & age_3gr = 1 & edu3grB = 2) w_psN = w_psN*563566.49980812/563566.500663863.
IF (gndr_r = 1 & age_3gr = 1 & edu3grB = 3) w_psN = w_psN*303931.972016129/303931.972447988.
IF (gndr_r = 1 & age_3gr = 2 & edu3grB = 1) w_psN = w_psN*524611.232753059/524611.231397038.
IF (gndr_r = 1 & age_3gr = 2 & edu3grB = 2) w_psN = w_psN*564319.670681312/564319.66970143.
IF (gndr_r = 1 & age_3gr = 2 & edu3grB = 3) w_psN = w_psN*469373.096838153/469373.098503725.
IF (gndr_r = 1 & age_3gr = 2 & edu3grB = 0) w_psN = w_psN*4866.16272525028/4866.16272905393.
IF (gndr_r = 1 & age_3gr = 3 & edu3grB = 1) w_psN = w_psN*767806.411263904/767806.409147767.
IF (gndr_r = 1 & age_3gr = 3 & edu3grB = 2) w_psN = w_psN*325888.813406563/325888.812429185.
IF (gndr_r = 1 & age_3gr = 3 & edu3grB = 3) w_psN = w_psN*255564.427733593/255564.428673281.
IF (gndr_r = 2 & age_3gr = 1 & edu3grB = 1) w_psN = w_psN*389035.693576249/389035.694779535.
IF (gndr_r = 2 & age_3gr = 1 & edu3grB = 2) w_psN = w_psN*515288.250780406/515288.250382046.
IF (gndr_r = 2 & age_3gr = 1 & edu3grB = 3) w_psN = w_psN*403802.605420875/403802.605834429.
IF (gndr_r = 2 & age_3gr = 2 & edu3grB = 1) w_psN = w_psN*491722.809602336/491722.810242967.
IF (gndr_r = 2 & age_3gr = 2 & edu3grB = 2) w_psN = w_psN*554526.195583982/554526.195097961.
IF (gndr_r = 2 & age_3gr = 2 & edu3grB = 3) w_psN = w_psN*485148.732964405/485148.734165687.
IF (gndr_r = 2 & age_3gr = 3 & edu3grB = 1) w_psN = w_psN*1122471.68605812/1122471.68357541.
IF (gndr_r = 2 & age_3gr = 3 & edu3grB = 2) w_psN = w_psN*325506.153850111/325506.15442477.
IF (gndr_r = 2 & age_3gr = 3 & edu3grB = 3) w_psN = w_psN*215511.289125417/215511.289352941.
EXECUTE .

```

\*region7.

```

IF (regionbe = 1) w_psN = w_psN*5102865.9875/5102865.98862934.
IF (regionbe = 2) w_psN = w_psN*835237.0075/835237.006453276.
IF (regionbe = 3) w_psN = w_psN*2811257.585/2811257.58491739.
EXECUTE .

```

```

COMPUTE w_ps.n = w_psN* 1798/8749360.58 .
EXECUTE .

```

DESCRIPTIVES  
VARIABLES=dweight w\_ps.n w\_psN  
/STATISTICS=MEAN STDDEV VARIANCE MIN MAX SEMEAN .