UNIVERZA V LJUBLJANI
FAKULTETA ZA DRUŽBENE VEDE

Tina Zupanič

Neodgovor na anketno vprašanje v raziskavah, povezanih z zdravjem

Magistrsko delo

Ljubljana, 2011

UNIVERZA V LJUBLJANI
FAKULTETA ZA DRUŽBENE VEDE

Tina Zupanič

Mentor: red. prof. dr. Vasja Vehovar

Neodgovor na anketno vprašanje v raziskavah, povezanih z zdravjem

Magistrsko delo

Ljubljana, 2011

Neodgovor na anketno vprašanje v raziskavah, povezanih z zdravjem

Pričujoče delo obravnava različne dejavnike, ki vplivajo na *neodgovor na anketo vprašanje*. Med dejavnike sodijo tematika ankete, značilnosti anketnih vprašanj, značilnosti anketiranca in anketarja kot tudi značilnosti načina anketiranja. V magistrskem delu uvodoma podamo širši teoretski pregled splošne problematike neodgovora spremenljivke kot tudi posebnosti na področju raziskav povezanih z zdravjem. V drugem empiričnem delu nato analiziramo zgoraj omenjene dejavnike v primeru dveh anket o zdravju, in sicer »Ankete o zdravju in zdravstvenem varstvu« (angl. »European Health Interview Survey«), izvedene leta 2007, ter ankete »Z zdravjem povezano vedenje v šolskem obdobju« (angl. »Health Behaviour in School-aged Children Survey«), izvedene leta 2010.

Ključne besede: anketna raziskava, neodgovor spremenljivke, neodgovor na anketno vprašanje, način anketiranja, zdravje

Item non-response in health related surveys

This master’s thesis addresses different factors that have an influence on item non-response within the surveys conducted. These factors include the following: survey question topics, a question structure, respondent and interviewers characteristics and survey mode attributes. The introductory part of the master's thesis presents a detailed theoretical overview about a phenomenon regarding the item non-response and its characteristics observed within health related surveys. The second part, which is of an empirical nature, brings about a discussion as a result of an empirical analysis of the above mentioned factors in a case of two health related surveys, the European Health Interview Survey 2007 and the Health Behaviour in School-aged Children Survey 2010.

Key words:survey, item non-response, survey mode, health

KAZALO

KAZALO TABEL IN SLIK 7

1 UVOD 10

1.1 Raziskovalna teza in relevantnost teme 13
1.2 Struktura naloge 17

2 NEODGOVORI V ANKETNIH RAZISKOVANJIH 20

2.1 Neodgovor spremenljivke 20
2.2 Dejavniki, ki vplivajo na pojav neodgovora spremenljivke 25

3 NEODGOVOR SPREMENLJIVKE V PRIMERU ANKET O ZDRAVJU 41

3.1 Problemi neodgovorov na področju zdravja 43

4 EMPIRIČNA ANALIZA NEODGOVORA SPREMENLJIVKE 53

4.1 Opis izbranih raziskav 53
4.2 Posebnosti anketiranja otrok in mladostnikov 55
4.3 Neodgovor spremenljivke med odraslo populacijo 63
4.4 Socio-demografske značilnosti nerespondentov 75
4.5 Reševanje problema neodgovorov 86

5 RAZPRAVA IN ZAKLJUČEK 109

6 LITERATURA 114

7 PRILOGA 126

7.1 Priloga 1 126
7.2 Priloga 2 127
7.3 Priloga 3 129

KAZALO TABEL IN SLIK

Tabela 3.1: Neodgovor spremenljivke in merjenje neenakosti v zdravju med mladimi. 50
Tabela 3.2: Neodgovor spremenljivke in merjenje neenakosti v zdravju med odraslo populacijo. 51
Tabela 4.1: Neodgovor spremenljivke na posamezna izbrana anketna vprašanja v HBSC – 1. del. 57
Tabela 4.2: Neodgovor spremenljivke na posamezna izbrana anketna vprašanja v HBSC – 2. del. 58
Tabela 4.3: Neodgovori spremenljivke v primeru vprašanj o splošnem zdravstvenem stanju v EHIS 2007. 66
Tabela 4.4: Neodgovori spremenljivke v primeru vprašanj o izbranih boleznih in bolezenskih stanjih v EHIS 2007. 67
Tabela 4.5: Neodgovori spremenljivke v primeru vprašanj pitju alkoholnih pijač in uživanju marihuane v EHIS 2007. 69
Tabela 4.6: Neodgovori spremenljivke v primeru vprašanj o vrstah dohodkov in skupnem dohodku gospodinjstva v EHIS 2007. 70
Tabela 4.7: Neodgovori spremenljivke v primeru izbranih vprašanj v EHIS 2007. 72
Tabela 4.8: Končni status ankete. 76
Tabela 4.9: Kakšno je vaše splošno zdravstveno stanje? Nerespondenti proti respondentom. 78
Tabela 4.10: Ali imate kakšno dolgotrajno bolezen ali dolgotrajno zdravstveno težavo? Nerespondenti proti respondentom. 79
Tabela 4.11: Oviranost zaradi zdravstvenih težav pri običajnih aktivnostih. Nerespondenti proti respondentom. 80
Tabela 4.12: Model logistične regresije (model 1). 85
Tabela 4.13: Izračun MSE v primeru izbranih spremenljivk. 100
Tabela 4.14: Število neodgovorov spremenljivke v primeru izbranih vprašanj in odgovorov zbranih s pomočjo anketarja. 105
Tabela 4.15: Število neodgovorov spremenljivke v primeru izbranih vprašanj in odgovorov zbranih s samoanketiranjem. 105
Tabela 4.16: Model logistične regresije (model 2). 108
Tabela 7.1: Kategorije izidov ankete. 126

Tabela 7.2: Izračun stopenj odgovora in stopenj sodelovanja v EHIS 2007.....126
Tabela 7.3: Izračun stopenj zavračanja in stopenj stika v EHIS 2007.....127
Tabela 7.4: Tip naselja. 127
Tabela 7.5: Spol x petletni starostni razredi. 128
Tabela 7.6: Statistična regija. 128
Tabela 7.7: Dosežena stopnja izobrazbe.....129
Tabela 7.8: Test multikolinearnosti - koeficienti.....129
Tabela 7.9: Test multikolinearnosti. 130

Slika 1.1: Odločevalski model neodgovora spremenljivke. 16
Slika 2.1: Stopnje neodgovorov na posamezna anketna vprašanja in stopnje neodgovorov na celotno anketo, SCA 1986–2005. 24
Slika 2.2: Kontinuitetni model sodelovanja v anketi. 25
Slika 2.3: Stopnje neodgovora spremenljivke na vprašanje o dohodku glede na leto izvedbe ankete. 36
Slika 2.4: Delež znižanja neodgovorov spremenljivke na vprašanje o dohodku zaradi uporabe tehnike »postopnega oženja izbire odgovorov«. 37
Slika 2.5: Model obdelave informacij v procesu odgovarjanja na anketna vprašanja. 40
Slika 3.1: Determinante zdravja..... 48
Slika 4.1: Splošen pregled neodgovora spremenljivke pri izbranih anketnih vprašanjih v HBSC. 58
Slika 4.2: Vzorec neodgovorov spremenljivke pri izbranih vprašanjih HBSC. 60
Slika 4.3: Najpogosteje pojavljajoči vzorci neodgovora spremenljivke v podatkih HBSC. 61
Slika 4.4: Najpogosteje pojavljajoči vzorci neodgovora spremenljivke v podatkih HBSC, brez najpogosteje ponavljajočega vzorca. 62
Slika 4.5: Splošen pregled neodgovora spremenljivke v EHIS. 73
Slika 4.6: Vzorec neodgovora spremenljivke v EHIS..... 74
Slika 4.7: Najpogosteje pojavljajoči vzorci neodgovora spremenljivke v podatkih EHIS. 75
Slika 4.8: Delež neodgovora enote po spolu med vsemi osebami zbranimi v vzorec. 77
Slika 4.9: Delež neodgovora enote po starostnih skupinah med vsemi osebami zbranimi v vzorec..... 77

Slika 4.10: Raziskovalni model socio-demografskih dejavnikov, za katere se predvideva, da vplivajo na pojav neodgovora spremenljivke. 84
Slika 4.11: Porazdelitev uteži zaradi neodgovora enote v EHIS. 99
Slika 4.12: Raziskovalni model vpliva zaznane ocenjene občutljivosti anketnega vprašanja in načina anketiranja na pojav neodgovora spremenljivke v EHIS. ..107

1 UVOD

Zbiranje podatkov o različnih populacijskih skupinah z anketnimi vprašalniki se je v zadnjem času zelo razširilo, tako s strani izvajalcev uradne statistike kakor tudi s strani različnih tržnih agencij. Zaradi »preobremenjenosti« prebivalcev z najrazličnejšimi anketami - kot tudi z neposrednim trženjem (kar mnogi včasih niti ne ločujemo) - ti vse pogosteje zavračajo sodelovanje, kar posledično vpliva tudi na kakovost izvedenih anket in izračunane ocene populacijskih parametrov.

Nesodelovanje oziroma neodgovarjanje izbranih oseb na anketo ali le na posamezna anketna vprašanja predstavlja enega izmed virov napak, ki se pojavljajo v anketnih raziskovanjih, in lahko pomembno vpliva na statistična sklepanja iz vzorca na populacijo, kadar neodgovori niso upoštevani pri nadaljnji obdelavi podatkov. Ločimo dve vrsti napak zaradi neodgovorov, in sicer se neodgovor pojavi takrat, ko izbrana oseba, ali ne želi odgovarjati na anketo, kar predstavlja *neodgovor enote* (angl. »unit non-response«) ali pa, kadar izbrana oseba ne želi odgovarjati na posamezna anketna vprašanja, kar predstavlja *neodgovor spremenljivke* (angl. »item non-response«) (Dillman in drugi 2002, 3–7; Bethlehem in Kersten 1985). Neodgovor spremenljivke oziroma neodgovor na anketno vprašanje predstavlja osnovni primer neodgovorov, pri čemer vrednost manjka pri ustrezni enoti, ki je bila sicer vključena v vzorec in smo jo uspeli anketirati, vendar odgovor pri posameznem vprašanju vseeno manjka zaradi zavrnitve odgovora, neveljavnega odgovora, nezmožnosti odgovora, nepazljivosti anketarja ipd. (Vehovar 1995). V Sloveniji je za podatke manjkajoče na ravni posamezne spremenljivke sicer uveljavljen pojem neodgovor spremenljivke, zato bomo v nadaljevanju besedila uporabljali pojem *neodgovor spremenljivke*, čeprav se v anketnih raziskovanjih o zdravju pogosto uporablja tudi pojem neodgovor na anketno vprašanje.

Na stopnje sodelovanja izbranih oseb vplivajo različni dejavniki, kot so tematika same ankete, tematika posameznih vprašanj, način anketiranja, vpliv anketarja idr. Glede na to, da je zdravje za večino posameznikov ena izmed pomembnejših, če ne celo najpomembnejša vrednota v življenju, anketna raziskovanja, povezana z zdravjem, ponavadi dosegajo relativno visoke stopnje sodelovanja, vendar pa se pogosto pojavljajo neodgovori spremenljivke, predvsem pri občutljivejših temah, kot so npr. vprašanja o različnih tveganih vedenjih (spolna vedenja, kajenje, uživanje alkohola, uživanje prepovedanih drog ipd.), vprašanja o dohodkovnem statusu, družinskih članih ipd.

Na stopnje neodgovora spremenljivke vplivajo značilnosti izbrane osebe (anketiranca) in anketarja, značilnosti vprašanj ter značilnosti načina anketiranja (Yan in Curtin 2010). Neodgovori spremenljivke se pojavijo zato, ker informacija bodisi ni na voljo, dana informacija ni uporabna ali pa, ker je dana uporabna informacija izgubljena. Vprašanje je lahko preskočeno namenoma ali zaradi napake, neznanega odgovora, nemožnega odgovora, je odgovor zunaj obsega možnih odgovorov, odgovora ni moč razvrstiti ali pa je neberljiv (de Leeuw 2001).

Vpliv značilnosti posameznih anketirancev obravnavajo različni avtorji, ki ugotavljajo, da imajo spol, starost, izobrazba, socialno-ekonomski položaj in druge značilnosti posameznika pomemben vpliv na pojav neodgovora spremenljivke.

Razmah različnih načinov anketiranja, do katerega je prišlo v zadnjih desetletjih, je povzročil, da imajo raziskovalci na voljo sicer več načinov anketiranja, uporabo katerih lahko prilagajajo posameznim populacijskim skupinam, vendar pa lahko napačen izbor načina anketiranja povzroči posledično večje napake v podatkih, kamor spada tudi neodgovor spremenljivke. Način anketiranja označujeta dve dimenziji, in sicer

sredstvo, s pomočjo katerega je anketiranje izvedeno (anketiranje na papirju, računalniško podprto anketiranje, telefonsko anketiranje, spletno anketiranje), ter vključenost anketarja v proces anketiranja (samoanketiranje proti anketiranju s pomočjo anketarja). Oboje lahko vpliva na pojav neodgovora spremenljivke, pri čemer se lahko v povezavi s tematiko samih anketnih raziskovanj problematika še stopnjuje. Vključenost anketarja v proces anketiranja lahko do neke mere znižuje stopnje neodgovora spremenljivke predvsem takrat, ko gre za relativno zapletena vprašanja s preskoki pri vprašalniku na papirju, ko npr. dobro usposobljen anketar na račun preskokov ne ustvarja neodgovora spremenljivke, medtem ko lahko za anketiranje neusposobljen posameznik pri samoanketiranju ob neupoštevanju navodil in preskokov posameznih vprašanj povzroči veliko količino neodgovorov spremenljivk. Ob primerjavi različnih načinov anketiranja se največje razlike med posameznimi načini pojavljajo, kadar gre za občutljivejša vprašanja. Pri občutljivejših vprašanjih namreč anketiranje s pomočjo anketarja povzroča večje deleže družbeno zaželenih odgovorov in več nedoslednih odgovorov pa tudi bolj natančne odgovore v primeru odprtih vprašanj, kar se izrazito pokaže tako pri osebnem kakor tudi telefonskem anketiranju (De Leeuw 2005).

Raziskave iz obdobja od 1992 do 2003 kažejo, da so stopnje neodgovora spremenljivke v primeru občutljivejših tematik višje pri računalniško podprtem telefonskem anketiranju kot pri anketiranju s pomočjo anketarja. Podobno se kaže tudi pri vprašanjih, povezanih s tveganimi vedenji, kjer so stopnje neodgovora spremenljivke višje pri podatkih, zbranih s samoanketiranjem, kot pri podatkih, zbranih s pomočjo telefonskega anketiranja (Alvik in drugi 2005). Pomemben vpliv na pojav neodgovora spremenljivke v raziskavah o vedenjih ima tudi odnos anketiranca do anketnih raziskovanj, pri čemer sta pomembna tako kakovost izkušenj z anketnimi raziskovanji kakor tudi količina izkušenj z njimi (Stocké 2006).

Osrednja tema magistrskega dela bo *neodgovor na anketno vprašanje* oz. *neodgovor spremenljivke*, in sicer bodo v magistrskem delu najprej obravnavani različni dejavniki, ki vplivajo na stopnje neodgovorov na anketna vprašanja, v nadaljevanju bo obravnavano tematsko področje v okviru anket o zdravju, kjer so stopnje neodgovorov na posamezna anketna vprašanja najbolj pereč problem, v zadnjem delu pa bo opravljena analiza neodgovorov na posamezna anketna vprašanja v primeru »Ankete o zdravju in zdravstvenem varstvu« (angl. »European Health Interview Survey 2007«), izvedene leta 2007, in ankete »Z zdravjem povezano vedenje v šolskem obdobju« (angl. »Health Behaviour in School-aged Children Survey«), izvedene leta 2010. Obe anketi sta bili izvedeni na osnovi mednarodno določene metodologije, zaradi česar so podatki mednarodno primerljivi, izvedel pa ju je Inštitut za varovanje zdravja.

1.1 Raziskovalna teza in relevantnost teme

Stopnje odgovorov v anketnih raziskovanjih so v razvitih državah v zadnjih treh desetletjih bistveno upadle (de Leeuw in de Heer 2002; Baruch in drugi 2008). Neodgovore lahko opazujemo na več ravneh, in sicer na makro ravni (npr. družba, kultura, ekonomska situacija), srednji ravni (npr. načrt raziskave) ter mikro ravni (npr. izbrana oseba, anketar) (de Leeuw in de Heer 2002, 42). Čeprav so za celostno razumevanje pojava neodgovorov pomembne vse tri ravni, bosta za potrebe analize manjkajočih odgovorov na anketna vprašanja v magistrski nalogi natančneje obravnavani srednja in mikro raven, in sicer predvsem z vidika izbrane osebe. V magistrskem delu bomo preverjali naslednjo hipotezo: *»Socio-demografski dejavniki vplivajo na neodgovore spremenljivke pogojno glede na tip oziroma tematiko anketnega vprašanja. Spol,*

starost, izobrazba, družbeni položaj anketiranca imajo pomemben vpliv na stopnje neodgovorov na različne tipe anketnih vprašanj (dejstvena, mnenjska, občutljiva vprašanja ...).«

Temeljno izhodišče za preverjanje zgornje hipoteze bo predstavljal odločevalski model neodgovora spremenljivke (Slika 1.1). Model namreč predstavlja štiri splošne razlage za neodgovore na posamezna anketna vprašanja, in sicer (Beatty in Herrmann 2002, 76):

a) neustreznost razumevanja zastavljenega vprašanja,
b) nizka motivacija.
c) odločitev za zadrževanje informacij, četudi izpolnjujejo predvidene cilje vprašanja, ali
d) prepričanje, da je neodgovor primernejši glede na natančnost postavljenega vprašanja.

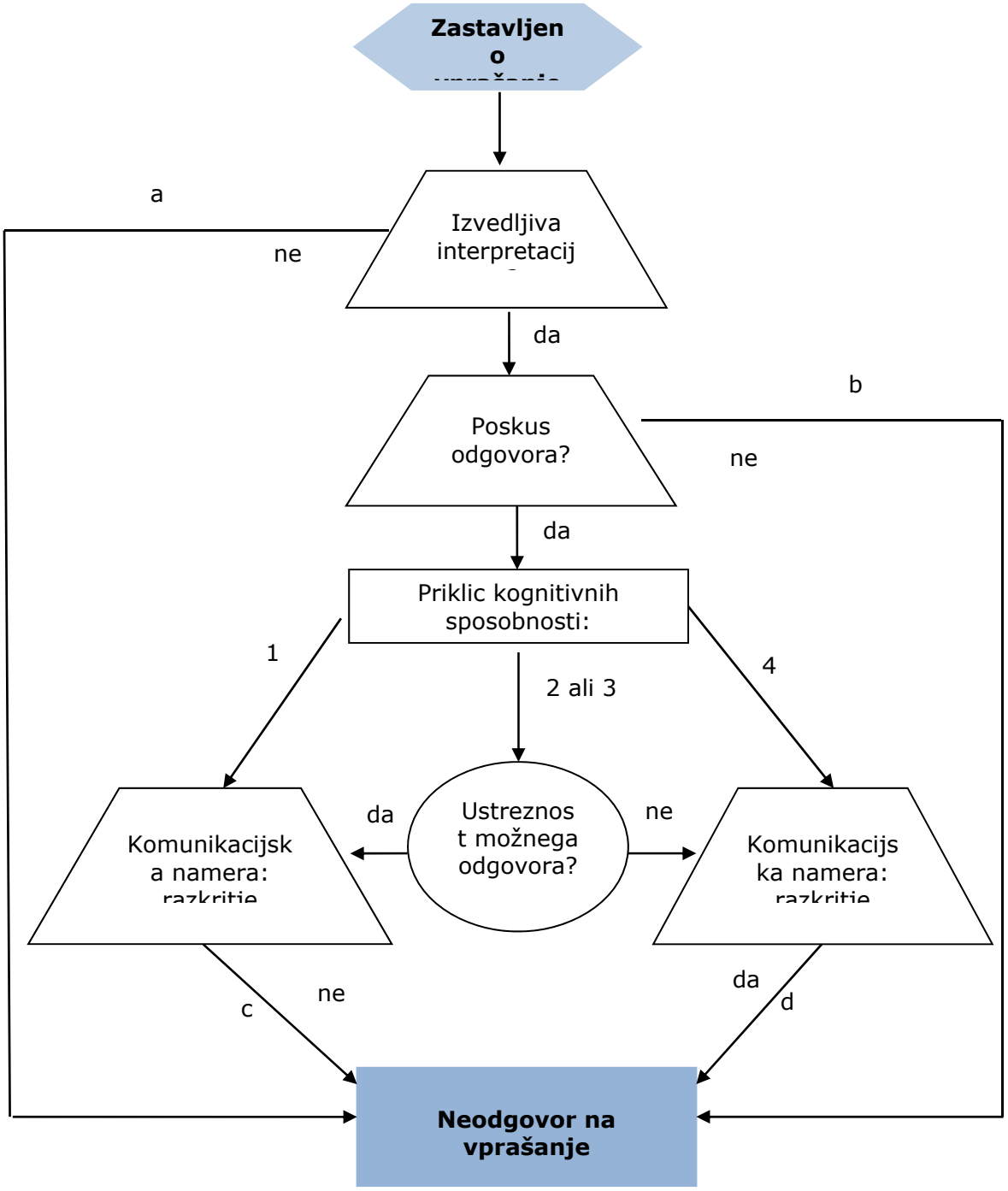
Razumevanje razlogov za neodgovore na anketna vprašanja nam lahko pomaga pri odločitvah glede oblikovanja strategij za zmanjšanje neodgovora spremenljivke. Proces odločanja se prične takrat, ko anketiranec prejme anketni vprašalnik. Na tem mestu mora anketiranec razumeti vprašanje in že tu se pojavi prva možnost neodgovora na vprašanje, če anketiranec ne razume vprašanja oziroma naloge, ki mu je bila z vprašanjem zastavljena. Proces odločanja se tu nadaljuje, če anketiranec zastavljeno vprašanje ustrezno razume in se odloči, ali bo odgovoril na zastavljeno vprašanje. Tu se prične anketirančeva prva odločitev glede na njegovo komunikacijsko namero: če je njegova motivacija nizka, lahko posameznik zaključi odločevalski proces. Odločitev za zaključek je lahko pogojena z dolžino anketnega vprašalnika in zahtevnostjo vprašanj. Če se odločevalski proces anketiranca nadaljuje, mora skozi priklic kognitivnih stanj priklicati relevanten in dostopen odgovor na zastavljeno vprašanje. Na tem mestu je možen vir neodgovora na vprašanje samo še komunikacijska namera anketiranca (kognitivno stanje 1). Kadar anketirancu odgovor na vprašanje ni znan (v primeru

kognitivnega stanja 4), je neodgovor na vprašanje najverjetnejši možen izid. Vendar pa odgovor anketiranca pogosto poteka v vmesnih stanjih, v katerih anketiranec ocenjuje kakovost možnega odgovora, saj pogosto ni očitno ali je dostopen odgovor tudi primeren odgovor na zastavljeno vprašanje. Anketiranci z istim razumevanjem vsebine vprašanja se lahko na podlagi razumevanja namena zastavljenega vprašanja ali odločijo odgovoriti ali pa na vprašanje ne odgovorijo. Na podlagi presoje anketiranci nadaljujejo z odločitvami komunikacijske namere, kar posledično vodi ali v odgovor ali pa v neodgovor na zastavljeno vprašanje (Beatty in Herrmann 2002, 76–77).

Iz spodaj predstavljenega odločevalskega modela neodgovora na anketno vprašanje lahko sklepamo, da imata pomemben vpliv na odločitev posameznika za neodgovor nezainteresiranost posameznika za odgovore ter nerazumevanje oziroma kompleksnost anketnega vprašanja. Predpostavljamo, da so mlajši anketiranci manj zainteresirani za korektno odgovarjanje na anketna vprašanja (predvsem na področju različnih tveganih vedenj, ki so družbeno morda bolj nesprejemljiva), medtem ko se pri starostnikih zaradi procesa staranja in težav s pomnjenjem lahko pojavljajo večji deleži manjkajočih odgovorov pri anketnih vprašanjih, kjer se morajo anketiranci za daljše časovno obdobje v preteklosti spomniti določenega dogodka.

Raziskave namreč ugotavljajo, da lahko značilnosti posameznih anketirancev pomembno vplivajo na deleže manjkajočih odgovorov. Deleži manjkajočih odgovorov sicer variirajo glede na posamezne populacije in tematike anketnih raziskovanj, vendar v splošnem velja, da so ti deleži višji pri starejših in nižje izobraženih posameznikih (Dillman in drugi 2002, 14).

Slika 1.1: Odločevalski model neodgovora spremenljivke.



Vir: Beatty in Herrmann 2002, 77.

Na deleže neodgovorov na anketno vprašanje poleg značilnosti posameznih anketirancev vplivajo tudi različni načini anketiranja. Preverjali bomo tudi naslednjo hipotezo, in sicer: *»Način anketiranja pomembno vpliva na neodgovore na anketno vprašanje. Samoanketiranje*

zvišuje verjetnost za pojav neodgovora spremenljivke, medtem ko je v primeru osebnega anketiranja s pomočjo anketarjev verjetnost za pojav neodgovora spremenljivke manjša.«

Različne raziskave namreč kažejo, da obstaja jasna povezava med načinom anketiranja in deležem neodgovora spremenljivke. Višji deleži neodgovorov spremenljivke se pojavljajo pri podatkih, zbranih s samoanketiranjem (npr. pri poštnih anketah) in samoanketiranjem z vprašalnikom na papirju. Samoanketiranje lahko izjemoma dosega višje deleže odgovorov na anketna vprašanja glede na ostale načine anketiranja pri občutljivejših vprašanjih (de Leeuw 2001, 150).

V splošnem velja, da se manjkajoči odgovori na anketna vprašanja pogosteje pojavljajo pri podatkih, zbranih s samoanketiranjem, kot pri podatkih, zbranih s pomočjo anketarjev (Tourangeau in drugi v Dillman in drugi 2002, 13). Pri samoanketiranju (kar še posebej velja za vprašalnike na papirju – PAPI) so odgovori na posamezna vprašanja odvisni predvsem od tega, ali anketiranec prebere vsako posamezno vprašanje, od vrstnega reda prebranih vprašanj in števila možnih odgovorov. Pomemben vpliv na manjkajoče odgovore na anketna vprašanja pri samoanketiranju imata tudi postavitev in izgled vprašalnika ter posameznih vprašanj (Jenkins in Dillman v Dillman in drugi 2002, 13).

1.2 Struktura naloge

V začetku prvega dela naloge bodo najprej obravnavane splošne značilnosti neodgovora spremenljivke ter povezanost neodgovora enote in neodgovora spremenljivke. V nadaljevanju bodo natančneje obravnavani dejavniki, ki vplivajo na pojav neodgovorov, pri čemer bo poseben poudarek na občutljivih anketnih vprašanjih in neodgovorih na vprašanja o

dohodku. Posebej pa bo obravnavana tudi ena izmed oblik neodgovora spremenljivke, in sicer odgovor »ne vem«.

V drugem delu naloge bodo identificirani problemi, ki jih povzroča neodgovor spremenljivke v okviru anket o zdravju in z zdravjem povezanih vedenjih, ter obravnavano tematsko področje v okviru anket, povezanih z zdravjem, kjer neodgovor spremenljivke raziskovalcem trenutno povzroča največje težave.

V tretjem delu naloge se bodo prepletale ugotovitve preteklih raziskav z analizami podatkov »Ankete o zdravju in zdravstvenem varstvu« (angl. »European Health Interview Survey 2007«), izvedene leta 2007, in ankete »Z zdravjem povezano vedenje v šolskem obdobju« (angl. »Health Behaviour in School-aged Children Survey«), izvedene leta 2010. Obe raziskavi bosta najprej na kratko opisani na začetku tega dela. V nadaljevanju bosta natančneje obravnavana neodgovor spremenljivke med mladimi in neodgovor spremenljivke med odraslo populacijo ter predstavljene socio-demografske značilnosti respondentov in nerespondentov v »Anketi o zdravju in zdravstvenem varstvu«. V primeru obeh anket bo najprej narejena analiza manjkajočih odgovorov na anketno vprašanje oz. »item non-response« (manjkajoči: »ne želim odgovoriti«, »ne vem«), pri čemer bodo natančneje analizirana tudi bolj »problematična« vprašanja, saj ima lahko tematika vprašanj pomemben vpliv na stopnje neodgovora spremenljivke. V primeru »Ankete o zdravju in zdravstvenem varstvu« bodo analizirane socio-demografske značilnosti posameznih skupin, ki niso odgovarjale na posamezna vprašanja, pri čemer bodo socio-demografske značilnosti oseb, ki na posamezna izbrana vprašanja niso odgovorili, primerjane tudi s socio-demografskimi značilnostmi drugih nerespondentov, ki so zavrnili sodelovanje v anketi. V primeru »Ankete o zdravju in zdravstvenem varstvu« bo s pomočjo modela logistične regresije izvedena tudi analiza vpliva posameznih socio-demografskih spremenljivk na pojav neodgovora spremenljivke. Nato

bodo obravnavane metode zmanjševanja vpliva neodgovorov, pri čemer bosta natančneje obravnavana uteževanje in vstavljanje manjkajočih podatkov. Uteževanje zaradi neodgovora enote bo zgolj za ilustracijo izvedeno v primeru »Ankete o zdravju in zdravstvenem varstvu«. V primeru »Ankete o zdravju in zdravstvenem varstvu« bo namreč dobljena vzorčna struktura, primerjana s populacijsko strukturo, na podlagi česar, bo nato izračunana utež zaradi neodgovora enote (izračun na podlagi populacijske strukture po spolu in starosti, izobrazbi, statistični regiji in tipu naselja). Pred uteževanjem in po njem bodo analizirana tudi tri kratka vprašanja o zdravstvenem stanju, ki so bila zastavljena tudi nerespondentom, na podlagi česar bo izračunano povečanje variance zaradi uteževanje in preverjena koristnost uteževanja. V drugem modelu bo s pomočjo logistične regresije natančneje obravnavan še vpliv načina anketiranja na neodgovor spremenljivke v »Anketi o zdravju in zdravstvenem varstvu«. Opravljena bo analiza vpliva načina anketiranja na neodgovor spremenljivke, pri čemer lahko znotraj »Ankete o zdravju in zdravstvenem varstvu« opazujemo dva različna načina anketiranja. Podatki v anketi so bili zbrani z vprašalnikom na papirju; večino vprašalnika je izvedel anketar, odgovori na vprašanja o samoplačniških stroških ter alkoholu, tobaku in drogah pa so bili zbrani s samoanketiranjem. Na drugi strani so bili vsi podatki, zbrani z anketo »Z zdravjem povezano vedenje v šolskem obdobju«, pridobljeni s samoanketiranjem otrok, starih 11, 13 in 15 let, z vprašalnikom na papirju.

Na koncu sledita še, razprava in zaključek, kjer bodo na podlagi rezultatov analize preverjene raziskovalne hipoteze in identificirana področja v okviru zdravstva, kjer lahko neodgovori na posamezna vprašanja pomembno izkrivljajo dejansko sliko stanja na tem področju.

2 NEODGOVORI V ANKETNIH RAZISKOVANJIH

Problem nepopolnih oziroma manjkajočih podatkov nastane v statistični analizi vsakič, ko manjka del podatkov, ki smo jih pričakovali. Neupoštevanje dejstva, da podatki manjkajo, lahko povzroči večje izkrivljanje rezultatov, posebej pri sklepanju na osnovi vzorca proučevane populacije. Običajna predpostavka verjetnostnih vzorcev, da so enote izbrane s pozitivno in vnaprej znano verjetnostjo, je namreč ogrožena, saj se enote z manjkajočimi podatki pogosto razlikujejo od drugih. Tako lahko npr. premožnejše osebe pogosteje zavračajo sodelovanje v anketi, kar vodi k podcenjenim ocenam o dohodkih (Vehovar 2007, 1).

2.1 Neodgovor spremenljivke

Ob pojavu neodgovora spremenljivke enota opazovanja, npr. izbrana oseba, sicer odgovarja na vprašanja, vendar so zaradi nekega razloga podatki v primeru določenih vprašanj nedostopni (de Leeuw 2001, 147). Neodgovor spremenljivke obsega naslednje možne izide (Kveder 2005):

a) zavrnitev odgovora (»ne želim odgovoriti«),
b) odgovor »ne vem« ter
c) preskok posameznega vprašanja.

Odločitev za štetje nekaterih odgovorov kot manjkajočih je odvisna od cilja raziskave. Tako je npr. odgovor »ne vem« pomemben odgovor na vprašanje o volilnih namerah anketiranca, medtem ko v primeru vprašanja o višini dohodka odgovor »ne vem« nima informacijske vrednosti in ga tako štejemo kot manjkajočega. Manjkajoč odgovor je tako definiran kot manjkajoč, če ga kot takega interpretira raziskovalec (de Leeuw 2001, 148). Neodgovore spremenljivke lahko razdelimo na tri glavne skupine (de Leeuw 2001, 149), in sicer:

a) sistematično manjkajoči podatki zaradi načrta raziskave,

b) vsi podatki so manjkajoči od neke točke anketnega vprašalnika dalje (delen neodgovor na anketo),
c) podatki manjkajo v primeru posameznih vprašanj in posameznih anketirancev (neodgovor na vprašanje).

Neodgovori spremenljivke se pojavijo ali zato, ker informacija ni na voljo ali dana informacija ni uporabna, ali pa zato, ker je dana uporabna informacija izgubljena. Vprašanje je lahko preskočeno iz več razlogov: namenoma ali pomotoma, ker je odgovor neznan, odgovor ni možen, odgovor je zunaj obsega možnih odgovorov, odgovora ni moč razvrstiti ali pa je neberljiv. Podatki so lahko manjkajoči na več načinov, in sicer manjkajoče vrednosti manjkajo popolnoma slučajno (MCAR – »missing compleatly at random«), manjkajoče vrednosti manjkajo slučajno (MAR – »missing at random«), ali pa manjkajoče vrednosti ne manjkajo slučajno (NMAR – »not missing at random«). Primer popolnoma slučajno manjkajočih vrednosti se pojavi takrat, ko anketiranec slučajno izpusti vprašanje, kadar pa so manjkajoči odgovori povezani z vrednostjo opazovanih podatkov, vendar ne z vrednostjo samega vprašanja, so podatki manjkajoči slučajno; npr. zaradi slabšega spomina se starejši anketiranec morda težje spomni določenega dogodka, pri čemer so manjkajoči odgovori povezani s starostjo in ne s samim dogodkom. Kadar pa so manjkajoči odgovori povezani z odgovorom na samo vprašanje, podatki ne manjkajo slučajno, npr. kadar anketiranec odgovori »ne vem«, zavrne odgovor, ali pa preskoči vprašanje zato, ker je njegov odgovor na zastavljeno vprašanje družbeno nesprejemljiv (npr. prekomerno pitje alkoholnih pijač) (de Leeuw 2001, 149-150; Widaman 2006).

V anketnih raziskovanjih sta v luči kakovosti samega raziskovanja in dobljenih populacijskih parametrov izredno pomembna pristranskost ocen zaradi neodgovorov in natančnost ocen populacijskih parametrov, vendar sta prepogosto zanemarjena s strani raziskovalcev, kar je še posebej izrazito v anketnih raziskovanjih med mladimi (Pike 2008).

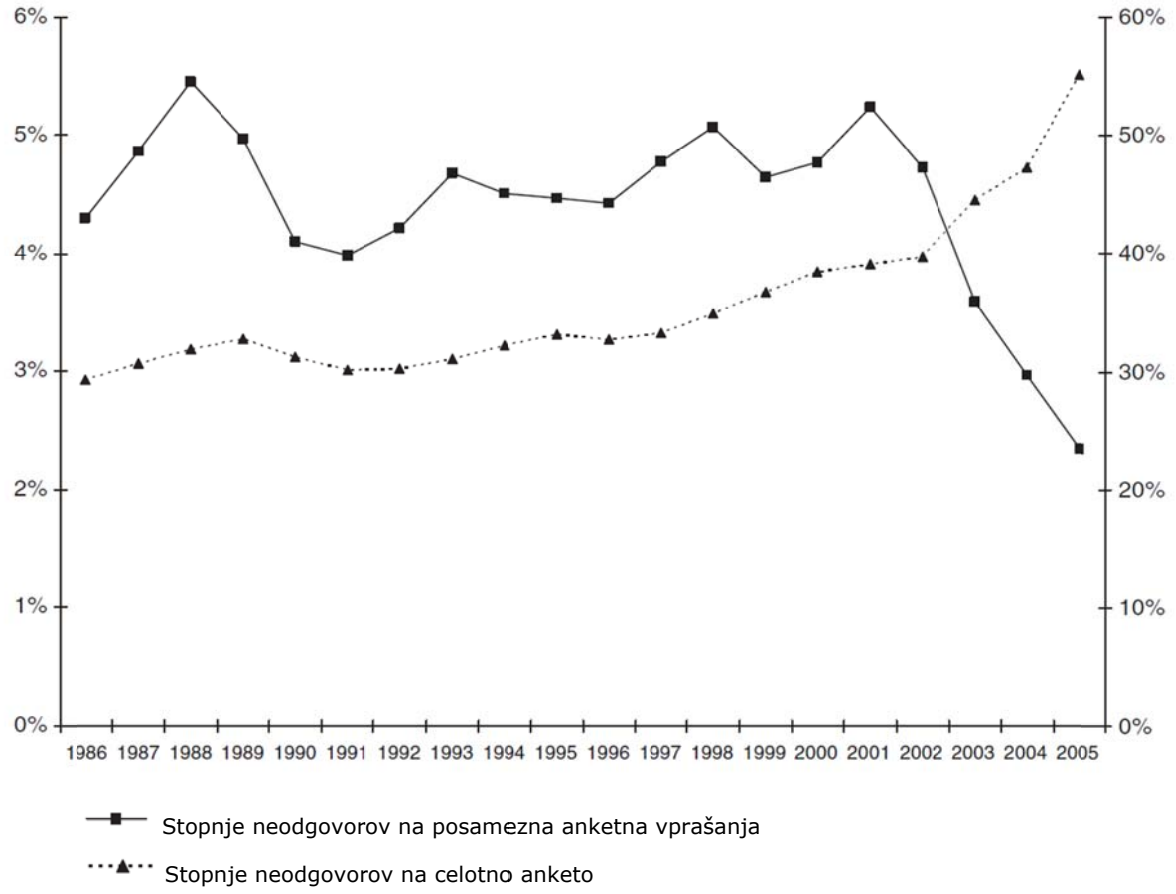
2.1.1 Povezanost neodgovora enote in neodgovora spremenljivke

Neodgovori predstavljajo velik problem za anketna raziskovanja, in sicer tako na ravni enote oziroma celotnega anketnega vprašalnika kakor tudi na ravni posameznega anketnega vprašanja. Tradicionalno se v anketnih raziskovanjih neodgovori enote in neodgovori na posamezna anketna vprašanja obravnavajo ločeno, z različnim vplivom na kakovost zbranih podatkov, različnimi statističnimi obdelavami in prilagoditvami ter nenazadnje z različnimi osnovnimi vzroki (Yan in Curtin 2010). V anketnem raziskovanju se neodgovore enote ponavadi obravnava kot večji problem, čeprav lahko tudi neodgovori spremenljivke pomembno vplivajo na kakovost izvedene ankete ter z njo dobljenih rezultatov. Ob upoštevanju samo tistih enot, ki so v celoti odgovorile na vsa anketna vprašanja, lahko neodgovor spremenljivke dodatno zmanjšuje velikost vzorca. Neodgovor spremenljivke lahko namreč predstavlja v anketnih raziskovanjih pomemben problem, kadar gre za nezanemarljiv delež manjkajočih podatkov. Nezanemarljivi manjkajoči podatki se pojavijo takrat, kadar je vzorec manjkajočih podatkov povezan z vrednostmi spremenljivke, ki nas zanima (Little in Rubin v Yan in Curtin 2010). Različni avtorji npr. Loosveldt in drugi (2002) ugotavljajo pozitivno povezanost med posameznikovo nagnjenostjo k sodelovanju v raziskavi in njegovo nagnjenostjo k odgovarjanju na posamezna anketna vprašanja. Posamezniki, ki sodelujejo v longitudinalnih raziskavah, pri katerih se pojavljajo visoki deleži neodgovorov spremenljivke, imajo večjo verjetnost, da v naslednjem valu raziskave ne bodo želeli več sodelovati, zato jim je treba posvetiti posebno pozornost ter jih poskusiti pritegniti k sodelovanju v naslednjem valu raziskave, npr. z dodatnimi spodbudami ali posebnimi predhodnimi pismi (Yan in Curtin 2010).

Anketna raziskovanja, ki imajo višje stopnje sodelovanja in torej nižje stopnje neodgovorov enote, imajo pogosto višje deleže neodgovorov spremenljivke, saj so med sodelujočimi pogosto večji deleži takšnih posameznikov, ki so sicer manj naklonjeni odgovarjanju na posamezna anketna vprašanja. Razumljivo je namreč, da se oseba, ki je izbrana v vzorec in jo anketar neprestano nagovarja k sodelovanju v raziskavi, sicer odloči za sodelovanje, vendar odgovori na čim manjše število vprašanj, kar posledično zvišuje stopnje neodgovora spremenljivke ter s tem zmanjšuje kakovost izvedene ankete in dobljenih populacijskih ocen (Yan in Curtin 2010).

Spodnja slika prikazuje trend letnega povprečnega indeksa deleža neodgovora spremenljivke v anketi potrošnikov, ki variira od 2,4 % do 5,5 % (zgornja krivulja, ki se nanaša na levo navpično os slike). Variiranja v deležu neodgovorov spremenljivke so nesistematična, pri čemer variacije z izjemo zadnjih nekaj let niso statistično značilno pomembne in so verjetno posledica razlik v motivaciji posameznikov, pripravljenih sodelovati v raziskavi, kar se kaže tako v stopnjah neodgovorov spremenljivke kakor tudi v stopnjah neodgovorov enote (Yan in Curtin 2010, 539). Trend upada neodgovorov spremenljivke lahko lažje razumemo z naraščajočim trendom neodgovorov enote (spodnja krivulja, ki se nanaša na desno navpično os slike). Trend stopenj neodgovorov namreč kaže premik stopenj neodgovorov spremenljivke skladno s premikom stopenj neodgovorov enote. Analiza, ki sta jo izvedla Yan in Curtin (2010), je namreč pokazala na statistično značilen upad stopenj neodgovorov spremenljivke ob povečanih stopnjah neodgovorov enote. Podatki namreč kažejo na pomembno negativen vpliv stopenj neodgovorov enote na stopnje neodgovorov spremenljivke.

Slika 2.1: Stopnje neodgovorov na posamezna anketna vprašanja in stopnje neodgovorov na celotno anketo, SCA 1986–2005.

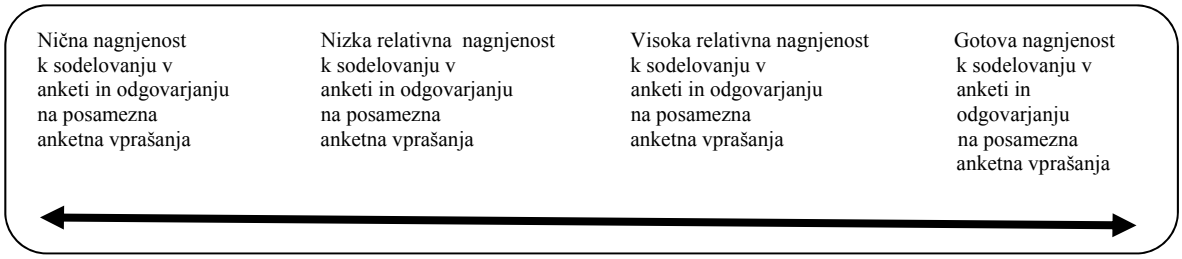


Vir: Yan in Curtin 2010, 540.

Kontinuitetni model sodelovanja v anketi (Slika 2.2) predpostavlja pozitivno povezanost med anketirančevo nagnjenostjo k neodgovoru enote in njegovim proizvajanjem neodgovorov spremenljivke. Yan in Curtin (2010) v analizi, ki sta jo izvedla na podlagi 20-letne serije indeksa neodgovorov spremenljivke, izbranih v anketah potrošnikov (SCA – »Survey of Consumers«), ugotavljata, da imajo anketiranci, ki so bili k sodelovanju v anketi spodbujeni z večkratnimi klici oz. stiki in prigovarjanjem, višje deleže neodgovorov spremenljivke kot anketiranci, ki so v sodelovanje v raziskavi privolili že ob prvem stiku. Tudi anketiranci, ki so prvotno sicer zavrnili sodelovanje v raziskavi in so se kasneje pustili prepričati v sodelovanje, imajo višje deleže neodgovorov spremenljivke kot anketiranci, ki so v sodelovanje privolili že ob prvem

stiku. Višji deleži neodgovorov spremenljivke se pojavljajo v primeru starejših anketirancev, pri ženskah, neporočenih, manj premožnih ter tistih, ki imajo srednjo šolo ali pa so nižje izobraženi (Yan in Curtin 2010; Riphahn in Seferling; de Leeuw 2001).

Slika 2.2: Kontinuitetni model sodelovanja v anketi.



Vir: Yan in Curtin 2010, 536.

2.2 Dejavniki, ki vplivajo na pojav neodgovora spremenljivke

Na stopnje neodgovorov spremenljivke vplivajo značilnosti izbrane osebe (anketiranca), značilnosti anketarja, značilnosti vprašanj ter značilnosti načina anketiranja. Tudi protokol raziskave in besedilne značilnosti posameznih vprašanj lahko pomembno vplivajo na stopnje neodgovora spremenljivke. Upad stopenj neodgovorov spremenljivke je povezan s spremembami merljivih značilnosti izbranih oseb (npr. stopnja izobrazbe, spol, višina dohodka) ter neizmerljivih značilnosti izbranih oseb, npr. občutljivost mejnih vrednosti, motiviranost za odgovarjanje na posamezna anketna vprašanja in splošna pripravljenost na sodelovanje v anketi (Yan in Curtin 2010).

Dillman in sodelavci (2002, 13–14) podobno ugotavljajo, da na pojav neodgovora spremenljivke vplivajo različni dejavniki, in sicer:

a) način anketiranja,
b) anketarji in njihovo izobraževanje,

c) tematika, struktura in zahtevnost vprašanj,
d) kontekstualni dejavniki ter
e) značilnosti izbranih oseb.

V splošnem velja, da se neodgovori spremenljivke pogosteje pojavljajo pri podatkih, zbranih s samoanketiranjem, kot pri podatkih, zbranih s pomočjo anketarjev (Tourangeau in drugi 2000). Pri samoanketiranju (kar še posebej velja za vprašalnike na papirju – PAPI) so odgovori na posamezna vprašanja vezani predvsem na to, ali anketiranec prebere vsako posamezno vprašanje, ter na vrstni red prebranih vprašanj in možnih odgovorov. Pomemben vpliv na neodgovor spremenljivke pri samoanketiranju imata tudi postavitev in izgled vprašalnika ter posameznih vprašanj (Jenkins in Dillman 1997).

Na zniževanje stopenj neodgovorov v anketnih raziskovanjih lahko pomembno vpliva tudi izobraževanje anketarjev. Anketarje se namreč na različne načine izobražuje o tem, kako se odzvati na posamezne načine neodgovora spremenljivke. Včasih se anketirance večkrat vpraša vprašanje na katerega ne odgovorijo in se jim ne ponuja možnosti »ne vem« ali »ne želim odgovoriti« (Dillman in drugi 2002, 13).

Pickery in Loosveldt (2001) ugotavljata, da je vpliv anketarja na pojav neodgovora spremenljivke povezan s težavnostjo samega vprašanja in s tem povezano kompleksnostjo procesa spraševanja, pri čemer je še posebej pomembna prisotnost t. i. filtrov.

Znano je, da se neodgovori spremenljivke pogosteje pojavljajo pri občutljivejših vprašanjih osebne narave, kot so npr. dohodek, spolno vedenje, pitje alkoholnih pijač, kršitve zakona (Tourangeau in drugi 2000).

Tudi struktura vprašanj lahko pomembno vpliva na deleže neodgovorov spremenljivke. Velja namreč, da so ti deleži večji pri samoanketiranju, odprtih in multiplih vprašanjih ter pri danih navodilih, kako odgovarjati na posamezna vprašanja, ki si jih mora anketiranec pri samoanketiranju predhodno prebrati (Featherston in Moy; Turner in drugi v Dillman in drugi 2002, 14).

Na večje deleže neodgovorov spremenljivke pogosto vpliva tudi zahtevnost posameznih vprašanj. Anketirancu lahko npr. predstavlja težavo vprašanje o tem, kolikokrat je obiskal zdravnika v zadnjih treh letih, saj se morda težko spomni, kolikokrat se je to dejansko zgodilo, medtem ko mu podobno vprašanje za obdobje zadnjih treh mesecev ali zadnji obisk ne predstavlja težav (Tomaskovic-Devey in drugi 1995).

Tudi značilnosti posameznih anketirancev lahko pomembno vplivajo na deleže neodgovorov spremenljivke. Deleži neodgovora spremenljivke sicer variirajo glede na posamezne populacije in tematike anketnih raziskovanj, vendar v splošnem velja, da so ti deleži višji v primeru starejših in nižje izobraženih posameznikov (Dillman in drugi 2002, 14). Tudi tematika posameznih vprašanj lahko pomembno vpliva na deleže neodgovorov spremenljivke. Kveder (2000) ugotavlja, da se vpliv spola na pojav neodgovora močno povezuje s tematiko vprašanj, verjetnost za pojav neodgovora pri moških v primeru vprašanj o rodnosti je namreč štirikrat višja kot pri ženskah.

Raziskave tudi kažejo različne stopnje neodgovora spremenljivke, vezano na posredne kazalnike kakovosti podatkov, in sicer se pojavljajo višji deleži neodgovorov spremenljivke med anketiranci, ki so sodelovali v pozni fazi izvedbe anketne raziskave, in tistimi, ki so prvotno sicer zavračali sodelovanje v anketi, kasneje pa vseeno privolili v sodelovanje (Fridman in drugi; Triplett in drugi; Willmack in drugi v Fricker in Tourangeau 2010). Z večanjem nagnjenosti k neodgovoru enote se

namreč višajo deleži neodgovorov na anketno vprašanje (Fricker in Tourangeau 2010).

Beatty in Herrman (2002) navajata tri dejavnike, ki pomembno prispevajo k stopnjam neodgovora spremenljivke, in sicer:

a) poznavanje tematike vprašanja s strani anketiranca,
b) presoja posameznikov o ustreznosti njihovega vedenja v primerjavi s stopnjo natančnosti, ki jo sámo vprašanje zahteva, ter
c) pripravljenost posameznikov za odgovarjanje na posamezna anketna vprašanja.

Prva dva dejavnika sta kognitivna, medtem ko je tretji dejavnik motivacijski. Zgornji model lahko uporabimo tudi v primeru vprašanj o dohodku, kjer se predvideva, da anketiranci ne odgovorijo na vprašanja o skupnem dohodku gospodinjstva, ker morda dejansko ne poznajo skupnega dohodka svoje družine, imajo morda le grobo oceno svojega skupnega dohodka, vendar predvidevajo, da morajo v odgovoru navesti natančno vsoto, ali pa enostavno svojega skupnega dohodka ne želijo povedati anketarju (Yan in drugi 2010).

Občutljiva vprašanja so tista vprašanja, ki so za posameznika občutljiva v smislu vsiljivosti oziroma se nanašajo na tabu teme, ki se zdijo posamezniku neprimerne v vsakodnevnih pogovorih ali pa so zunaj meja pogovora s priložnostnimi znanci kot so npr. anketarji. Zaznana občutljivost je določena predvsem z vsebino vprašanja, ki je označena kot vdor v zasebnost (Tourengau in drugi 2000; Tourengau in Yan 2007).

Beebe in sodelavci (2008) ugotavljajo, da anketiranci želijo sodelovati v raziskavi, kadar jih tematika ankete zanima in kadar poznajo organizacijo, ki raziskavo izvaja (višje stopnje sodelovanja namreč dosegajo anketna raziskovanja, ki jih izvajajo vladne organizacije).

Občutljiva anketna vprašanja

Postavljanje občutljivih vprašanj v anketnem raziskovanju je lahko izredno problematično, saj je nagnjeno k napakam in pristranskosti (Barnett 1998).

Občutljiva anketna vprašanja so npr. vprašanja o uporabi drog, spolnem vedenju, zdravstvenem stanju, glasovanju in dohodku. Pri teh vprašanjih se ponavadi pojavljajo tudi večje stopnje neodgovorov spremenljivke in večje merske napake kot v primeru drugih manj občutljivih anketnih vprašanj (Tourangeau in Yan 2007).

Občutljiva so tista anketna vprašanja, ki jih anketiranec zaznava kot *vsiljiva* in pri katerih obstaja *nevarnost razkritja informacij*. Vsiljiva anketna vprašanja ponavadi zadevajo tabu teme in teme, ki so označene kot neprimerne v vsakodnevnih pogovorih. Anketirancu pomenijo vdor v njegovo zasebnost, ne glede na to, kakšen je za anketiranca pravilen odgovor. Pomen občutljivosti posameznega anketnega vprašanja je bolj kot s situacijskimi dejavniki, kot sta npr., kje je zastavljeno določeno vprašanje in komu je zastavljeno, določen z vsebino vprašanja. V to kategorijo sodijo npr. vprašanja o posameznikovem dohodku in verskem prepričanju. Posamezniki namreč pogosto menijo, da se določena vsebinska področja enostavno ne tičejo anketarjev. Z vprašanji, ki spadajo v to skupino, pogosto tvegamo, da užalimo vse anketirance, ne glede na njihov položaj v okviru posameznega anketnega vprašanja. Na drugi strani pa nevarnost razkritja informacij pomeni skrb glede možnih posledic v primeru razkritja resničnega odgovora tretjim osebam. Anketno vprašanje je v tem smislu občutljivo, kadar zbuja strah glede posledic možnega razkritja odgovorov institucijam ali posameznikom, ki niso neposredno povezani z anketno raziskavo. Skrb anketirancev glede zaupnosti njihovih odgovorov pogosto variira in čeprav anketne raziskave anketirancem zagotavljajo zaupnost in nerazkritje, anketiranci v to

pogosto ne verjamejo, zato so skrbi glede razkritja informacij pomemben dejavnik napačnega poročanja o nelegalnih ali družbeno nesprejemljivih vedenjih (Tourangeau in Yan 2007).

Občutljivost posameznega anketnega vprašanja je močno povezana tudi s *konceptom družbene zaželenosti*, in sicer glede na to, v kakšnem obsegu vprašanje izzove odgovore, ki so družbeno nesprejemljivi in družbeno nezaželeni (Tourangeau in drugi 2000). Na eni strani so odgovori na posamezna anketna vprašanja glede vedenj in odnosov, ki ustrezajo družbenim normam in se zdijo družbeno zaželeni, na drugi strani pa so tisti odgovori, ki pomenijo deviantnost oziroma odklonskost v primerjavi z družbenimi normami in so označeni kot družbeno nezaželeni. Posamezno anketno vprašanje je občutljivo, kadar sprašuje o družbeno nezaželenem odgovoru, torej kadar posameznik z odgovarjanjem nanj prizna, da je kršil družbeno normo. Občutljivost je v tem primeru določena z anketirančevim potencialnim odgovorom na anketno vprašanje (Tourangeau in Yan 2007).

Občutljiva vprašanja vplivajo na tri anketne izide, in sicer na (Tourangeau in Yan 2007, 862):

a) stopnje odgovorov enote (ang. »unit response rate«),
b) stopnje neodgovora spremenljivke (angl. »item nonresponse rate«) ter
c) točnost odgovorov, ki pomeni delež anketirancev, ki so po resnici odgovorili na posamezno anketno vprašanje.

Občutljiva vprašanja povzročajo probleme v vseh treh zgoraj navedenih primerih, saj znižujejo tako splošne stopnje odgovorov enote, kakor tudi stopnje odgovorov spremenljivke in točnost odgovorov. Slednje predstavlja sistematično nenatančno poročanje v primeru občutljivejših vprašanj, saj anketiranci pogosto dosledno podporočajo nekatera npr. družbeno nezaželena vedenja, na drugi strani pa dosledno nadporočajo nekatera npr. družbeno zaželena vedenja, kar lahko pripomore k večji

pristranskosti vzorčnih ocen. Anketiranci pogosto podporočajo uporabo prepovedanih drog, uporabo alkohola, podatke o splavih, bankrotih, rabi energije, določenih virih dohodka in kriminalnih vedenjih (Tourangeau in Yan 2007).

V Sloveniji Zakon o varstvu osebnih podatkov v 6. členu opredeljuje *občutljive osebne podatke* kot »podatke o rasnem, narodnem ali narodnostnem poreklu, političnem, verskem ali filozofskem prepričanju, članstvu v sindikatu, zdravstvenem stanju, spolnem življenju, vpisu ali izbrisu v ali iz kazenske evidence ali evidenc, ki se vodijo na podlagi zakona, ki ureja prekrške; občutljivi osebni podatki so tudi biometrične značilnosti, če je z njihovo uporabo mogoče določiti posameznika v zvezi s kakšno od prej navedenih okoliščin« (ZVOP-1-UPB1).

Ker tudi podatki o zdravstvenem stanju prebivalcev spadajo med občutljive osebne podatke, je treba z vidika varovanja osebnih podatkov z njimi ravnati posebej pazljivo, zaradi svoje občutljivosti pa tudi pri posameznikih povzročajo, da v primeru zastavljenih vprašanj v anketnih raziskovanjih nanje pogosteje ne želijo odgovoriti.

Občutljiva anketna vprašanja namreč pogosto sprašujejo o družbeno nezaželenih oz. nesprejemljivih vedenjih ali osebnih zadevah posameznikov, o katerih izbrane osebe ne želijo razpravljati z neznanci, kot so npr. anketarji. Ker lahko takšna tematika vprašanj pri anketirancu izzove strah, lahko poveča tudi stopnje neodgovorov spremenljivke (Bradburn in drugi 1978). Vendar pa stopnja odgovorov »ne vem« ni povezana z občutljivostjo vprašanj, saj odgovor »ne vem« ne predstavlja prijaznejšega zavračanja odgovora. Anketiranci npr. pogosto zavračajo odgovarjanje na vprašanje o dohodku, na vprašanja o svojih vedenjih in osebnih značilnostih pa anketiranci ne odgovarjajo z »ne vem«, saj svoja vedenja in značilnosti zelo dobro poznajo, vendar jih le ne želijo poročati (Shoemaker in drugi 2000).

Ob velikem številu raziskav, ki anketirance sprašujejo o izredno občutljivih temah, kot je npr. uporaba prepovedanih drog, se pogosto sprašujemo o tem, ali anketiranec na vprašanje dejansko odgovori po resnici. Metodološke raziskave o točnosti poročanja v anketah o uporabi prepovedanih drog in drugih občutljivih temah ugotavljajo, da velik vir napake in pristranskosti ocen predstavlja napačno poročanje. Družbeno zaželene odgovore v primeru občutljivih vprašanj lahko zmanjšamo npr. s samoanketiranjem, med mladimi pa tudi z anketiranjem brez prisotnosti staršev (Tourangeau in Yan 2007).

Anketiranec lahko v primeru občutljivega vprašanja, npr. o uporabi prepovedanih drog v zadnjih dvanajstih mesecih, izbere naslednje načine odgovora (Tourangeau in Yan 2007, 877):

a) pove resničen odgovor,
b) neresnično poroča o pogostosti ali količini uporabe prepovedanih drog,
c) neresnično poroča, da v celoti zavrača uporabo prepovedanih drog, ali
d) zavrne odgovor na to vprašanje.

Anketiranci zavrnejo odgovor na posamezno vprašanje, kadar ne poznajo odgovora, kadar poznajo približen odgovor, vendar menijo, da morajo odgovoriti z natančnim odgovorom, ali pa, kadar enostavno ne želijo odgovoriti na zastavljeno vprašanje (Tourangeau in Yan 2007).

Na zastavljena občutljiva vprašanja anketiranci pogosto raje odgovorijo neresnično, kakor da zavrnejo odgovor na izbrano vprašanje, saj je zavračanje odgovora lahko razkrivajoče samo po sebi, npr. če nekdo zavrne odgovor na vprašanje, ali je že kdaj užival kokain, lahko to razkriva pritrdilen odgovor (Schaeffer 2000).

Občutljiva vprašanja predstavljajo širok spekter vprašanj, ki ne obsegajo samo tistih vprašanj, ki sprožajo vprašanje družbene zaželenosti, ampak tudi tista vprašanja, ki se anketirancu zdijo vsiljiva ali pa zbujajo skrb glede možnih posledic v primeru razkritja informacij. Posedovanje kokaina npr. ni samo družbeno nesprejemljivo, ampak je tudi nelegalno, zato lahko posamezniki v anketi o drogah ne poročajo o tem, saj menijo, da se bodo s tem izognili pravnim posledicam, in ne zgolj ustvarili primeren vtis (Tourangeau in Yan 2007).

Neodgovori na vprašanja o dohodku

Anketni podatki o osebnem dohodku in dohodku gospodinjstva so pogosto povezani z velikimi deleži manjkajočih odgovorov. Točni in popolni podatki o osebnem dohodku in dohodku gospodinjstva so izrednega pomena za uradno statistiko, vladne politike ter družbene vede in raziskovanje. Kadar so podatki o dohodku zbrani s pomočjo anketnih raziskav, so pogosto povezani z visokimi deleži manjkajočih podatkov. Različne raziskave ugotavljajo, da je stopnja neodgovorov na vprašanja o dohodku med 20 % in 40 %, medtem ko se stopnja neodgovorov na druga anketna vprašanja giblje med 1 % in 4 % (De Leeuw 1992; Moore in drugi 1999; Juster in Smith 1997).

Ker so anketna vprašanja o dohodku z vidika analize neenakosti v zdravju (angl. »Health Inequalities«) izredno pomembna, saj predstavljajo enega izmed kazalnikov posameznikovega socio-ekonomskega položaja, v nadaljevanju problem pojava neodgovora spremenljivke na vprašanje o dohodku obravnavamo bolj natančno, saj nas zanimajo vzroki za ta pojav ter možne rešitve za zmanjšanje neodgovora spremenljivke na vprašanje o dohodku.

Vprašanja o dohodku spadajo v kategorijo vprašanj, ki so za posameznika občutljiva v smislu vsiljivosti. Anketiranci namreč pogosto zaznavajo

takšna vprašanja kot osebno stvar, ki ne zadeva anketarjev (Yan in drugi 2010). Podobno tudi Beatty in Herrmann (2002) ugotavljata, da so neodgovori spremenljivke na vprašanja o dohodku zaradi svoje občutljivosti predvsem vir anketirančeve motivacije, njun model pa predvideva, da se neodgovori na vprašanja o dohodku pojavljajo v vseh populacijskih podskupinah.

Yan in sodelavci (2010) ugotavljajo, da se deleži neodgovorov spremenljivke na vprašanje o dohodku v opazovanem daljšem časovnem obdobju znižujejo, znižanje pa je povezano z anketiranci, ki so v splošnem pripravljeni sodelovati v anketi in odgovarjati na vsa anketna vprašanja.

Raziskave kažejo (npr. Körmendi; Schräpler; De Leeuw), da različni načini anketiranja vplivajo na deleže neodgovorov spremenljivke na vprašanja o dohodku. Schräpler in drugi (2006) npr. ugotavljajo, da se pri anketiranju s pomočjo računalnika (CAPI – angl. »Computer-Assisted Personal Interviews«) v primerjavi z anketiranjem na papirju (PAPI – angl. »Paper and Pencil Personal Interviews«) pojavljajo večji deleži neodgovorov spremenljivke na vprašanja o dohodku, medtem ko poštne ankete v primerjavi z anketiranjem s pomočjo računalnika ne izstopajo. Na drugi strani pa De Leeuw (2001) ugotavlja, da se večji deleži neodgovorov spremenljivke na vprašanja o dohodku pojavljajo pri podatkih, zbranih s samoanketiranjem, manjši pa pri podatkih, zbranih s pomočjo anketarja.

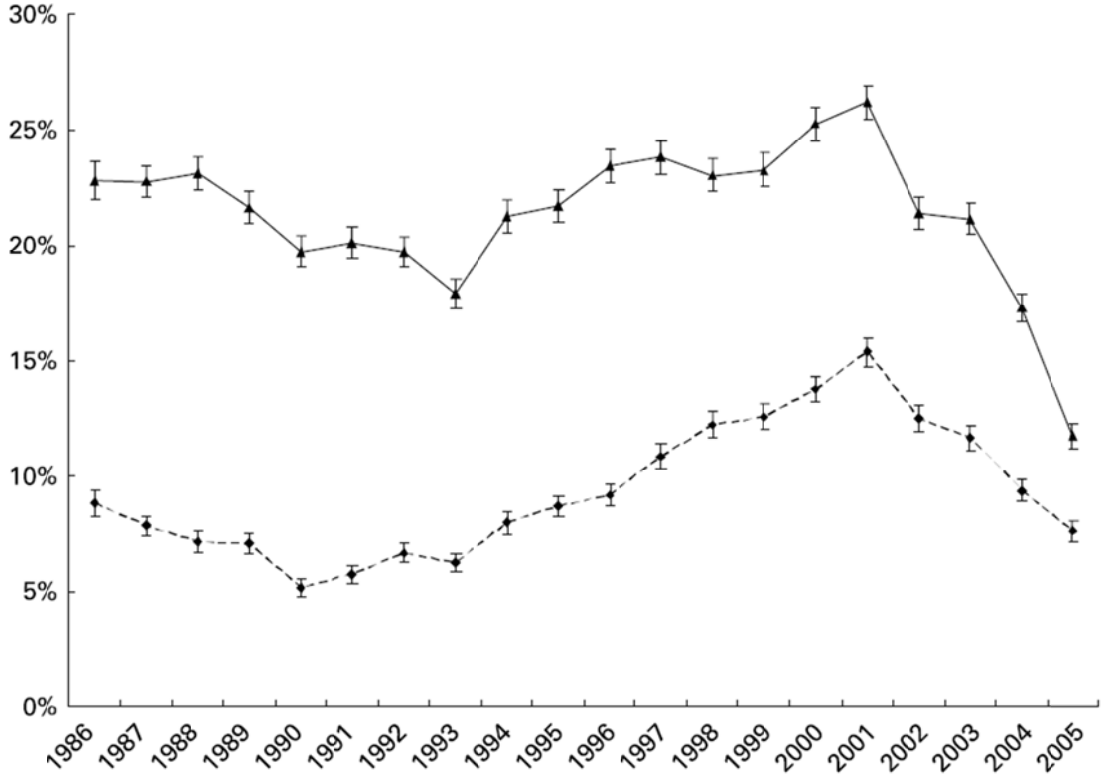
Tudi oblika vprašanj o dohodku pomembno vpliva na deleže neodgovorov. Vpliv oblike vprašanj o dohodku na neodgovor spremenljivke so obravnavali različni avtorji (Heeringa in drugi; Juster in Smith), in sicer so primerjali stopnje neodgovora spremenljivke na odprta vprašanja (angl. »open-ended questions«), ki anketiranca neposredno sprašujejo o dejanski, natančni vrednosti njegovega dohodka, in stopnje neodgovora spremenljivke na intervalna vprašanja o dohodkovnem razredu, v katerega se uvršča posamezen anketiranec (Yan in drugi 2010). Slednji

način predstavlja t. i. tehniko »postopnega oženja izbire odgovorov« (angl. »unfolding bracket«), s pomočjo katere zmanjšujemo stopnje neodgovora spremenljivke na vprašanja o dohodku (Heeringa in drugi 1995; Juster and Smith 1997; Vazquez Alvarez 2001; Winter 2002). S tehniko »postopnega oženja izbire odgovorov« nerespondentom na odprto vprašanje o dohodku, včasih pa kar vsem anketirancem, ponudimo vprašanja, ki anketirancu omogočajo, da se ali uvrsti v določen dohodkovni razred ali pa ga vprašamo o okvirni višini dohodka, npr. »Ali je skupen dohodek vašega gospodinjstva višji od 1000 EUR?«, in nato, če anketiranec odgovori z »da«, nadaljujemo npr. z vprašanjem, »Ali je višji od 1500 EUR?«, če pa z »ne« pa z vprašanjem, »Ali je nižji od 500 EUR?« Ta tehnika raziskovalcem omogoča zbiranje delnih informacij o anketirančevem dohodku, četudi anketiranec zavrača odgovarjanje na vprašanje o točni vrednosti svojega dohodka oz. dohodka gospodinjstva. Različne raziskave ugotavljajo, da tehnika »postopnega oženja izbire odgovorov« zmanjšuje deleže neodgovorov na vprašanja o dohodku tudi za več kot 50 %. Raziskave tudi ugotavljajo, da je tehnika »postopnega oženja izbire odgovorov« učinkovitejša v primeru kognitivnih in informacijskih dejavnikov, ki povzročajo neodgovore na vprašanja o dohodku, manj pa v primeru motivacijskih dejavnikov. Ta tehnika učinkovito zmanjšuje deleže neodgovora spremenljivke na vprašanja o dohodku, saj z nudenjem različnih fiksnih kategorij dohodka zmanjšuje kognitivno breme anketnega vprašanja na anketirance ter s tem zmanjšuje zaznano občutljivost vprašanja s strani anketiranca. Tehnika »postopnega oženja izbire odgovorov« je manj učinkovita pri anketirancih, ki ne želijo odgovoriti na vprašanje o dohodku, bistveno bolj pa pri anketirancih, ki so na osnovno odprto vprašanje o dohodku odgovorili z »ne vem« (Yan in drugi 2010; Juster in drugi 2006).

Neodgovor na vprašanja o dohodku je nezanemarljiv (angl. »nonignorable«), kar pomeni, da vzorec manjkajočih podatkov korelira znotraj vrednosti spremenljivke, ki nas zanima, v tem primeru dohodka

(Yan in drugi 2010). Raziskave (Greenless in drugi 1982; Lillard in drugi 1986; Riphahn in Serfling 2005) ugotavljajo pogojno povezanost deleža neodgovora spremenljivke na vprašanja o dohodku z višino anketirančevega dohodka, in sicer je delež neodgovorov spremenljivke na vprašanja o dohodku višji v primeru anketirancev z višjimi dohodki.

Slika 2.3: Stopnje neodgovora spremenljivke na vprašanje o dohodku glede na leto izvedbe ankete.



Stopnje neodgovorov spremenljivke na odprto vprašanje o skupnem dohodku gospodinjstva

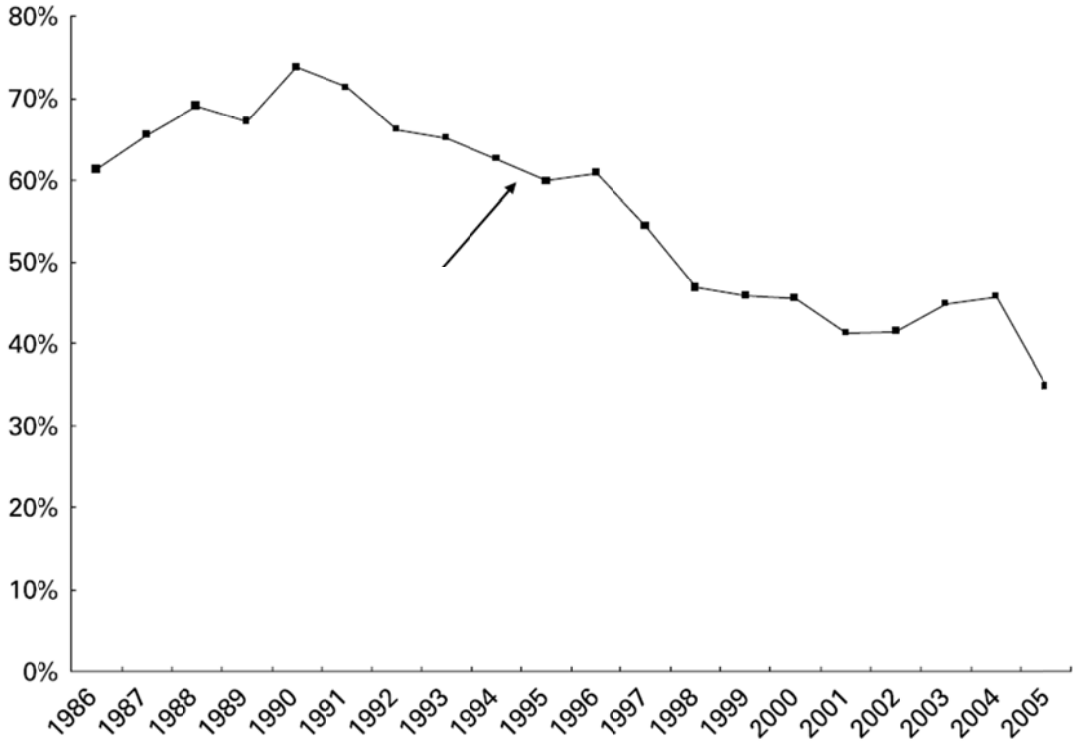
Stopnje neodgovorov spremenljivke na vprašanje o skupnem dohodku gospodinjstva ob uporabi tehnike »postopnega oženja izbire odgovorov«

Vir: Yan in drugi 2010, 151.

Zgornja slika (Slika 2.3) prikazuje stopnje neodgovora spremenljivke na vprašanje o dohodku glede na leto izvedbe ankete, pri čemer zgornja krivulja prikazuje stopnje neodgovorov spremenljivke na odprto osnovno vprašanje o skupnem dohodku gospodinjstva, spodnja krivulja pa stopnje neodgovorov spremenljivke na vprašanje o skupnem dohodku gospodinjstva ob uporabi tehnike »postopnega oženja izbire odgovorov«.

Yan in sodelavci (2010) ugotavljajo, da tehnika »postopnega oženja izbire odgovorov« učinkovito znižuje delež neodgovorov spremenljivke v primerjavi z deležem neodgovorov spremenljivke na odprto osnovno vprašanje o skupnem dohodku gospodinjstva, kar je razvidno tudi s slike na naslednji strani (Slika 2.4).

Slika 2.4: Delež znižanja neodgovorov spremenljivke na vprašanje o dohodku zaradi uporabe tehnike »postopnega oženja izbire odgovorov«.



Vir: Yan in drugi 2010, 152.

Na deleže neodgovora spremenljivke na vprašanje o dohodku podobno kot na neodgovore spremenljivke v primeru ostalih anketnih vprašanj vplivajo tudi starost, spol in izobrazba anketiranca. Za razumevanje visokih stopenj neodgovorov na občutljiva vprašanja, kot je npr. dohodek anketiranca, moramo razumeti determinante vedenja povezanega z neodgovori. Deleži neodgovorov spremenljivke v primeru vprašanj o dohodku posameznika ali gospodinjstva so namreč zgoščeni na repih porazdelitve dohodka. Zavračanje odgovarjanja na posamezna anketna

vprašanja o dohodku je povezano tudi s tem, ali je oseba zaposlena za polni ali skrajšani delovni čas, kar lahko razlagamo s pomočjo modela racionalne izbire, in sicer če so dohodki anketirancev zaposlenih za skrajšani delovni čas relativno nizki in želijo ti anketiranci predstaviti svoj osebni uspeh na trgu delovne sile, bodo verjetno zavrnili odgovor na vprašanje o dohodku. Večji deleži neodgovorov spremenljivke se pojavljajo tudi pri starejših anketirancih, pri samoanketiranju, če je anketar ženskega spola, kar je še posebej izrazito takrat, kadar je tudi izbrana oseba ženska ter med višje izobraženimi, medtem ko so deleži neodgovorov glede na velikost gospodinjstva, v katerem posameznik živi, nižji v primeru večjih gospodinjstev (Riphahn in Serfling 2005).

V anketah o zdravju in vedenjih, povezanih z zdravjem, je vprašanje o skupnem dohodku gospodinjstva ponavadi umeščeno za vprašanje o posameznih virih dohodkov in prejemkov članov gospodinjstva. Slednje naj bi namreč vplivalo na točnost poročane skupne vsote dohodka gospodinjstva (Daveren in drugi 2005).

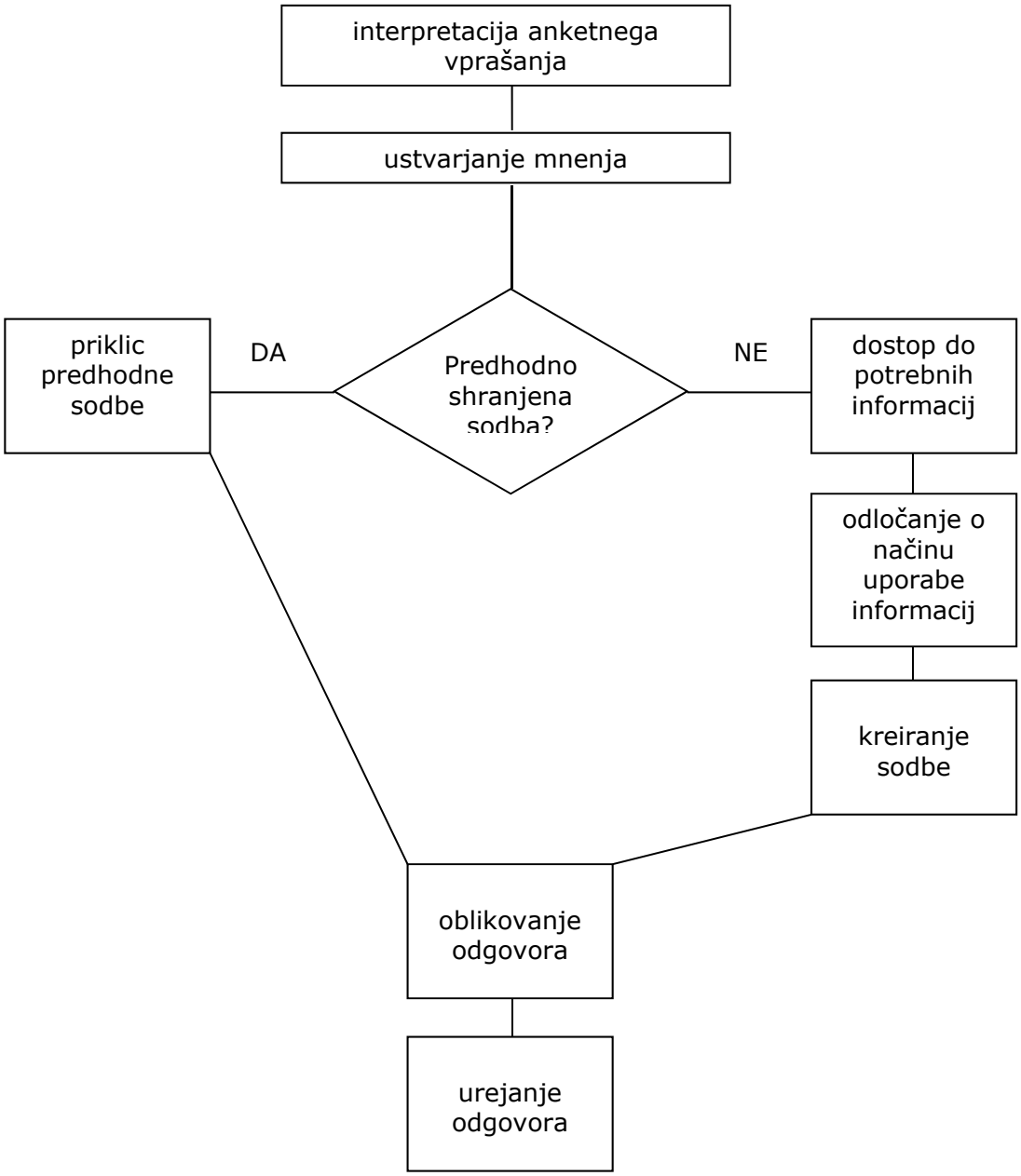
Odgovor »ne vem«

Odgovor »ne vem« predstavlja eno izmed oblik neodgovora spremenljivke in lahko poleg dejanskih zavrnitev odgovora na anketno vprašanje predstavlja pomemben delež neodgovora spremenljivke v anketnih raziskovanjih, predvsem ko gre za občutljivejša anketna vprašanja.

Raziskave (Juster in Smith; Riphahn in Serfling) ugotavljajo, da je treba odgovore »ne vem« in tiste, ki dejansko zavračajo odgovor na posamezno vprašanje (»ne želim odgovoriti«) obravnavati različno, kadar izvajamo vstavljanje oz. imputiranje manjkajočih podatkov, saj moramo gledati na odgovore »ne vem« kot na samostojne izide in jih ne smemo mešati s tistimi, ki dejansko ne želijo odgovoriti.

Shoemaker in sodelavci (2000) v svoji študiji ugotavljajo, da se večji deleži zavrnitev odgovorov pojavljajo pri občutljivejših vprašanjih, medtem ko se odgovori »ne vem« v večjih deležih pojavljajo pri vprašanjih, pri katerih je za odgovor potreben večji kognitivni napor. Anketiranci namreč pogosto dejansko ali nimajo mnenja ali pa védenja o tematiki, o kateri jih sprašuje anketar. Pri takšnih vprašanjih gre za to, da je za odgovor nanje potreben visok kognitivni napor, ki je za nekatere dele populacije preveč zahteven. Kompleksnost procesa odgovarjanja na anketna vprašanja so proučevali Sudman in sodelavci (1996). Proces vključuje interpretacijo vprašanja, pridobivanje informacij, ustvarjanje mnenja oziroma predstavitev ustreznega vedenja ter oblikovanje in urejanje odgovora. Pri težavnejših vprašanjih nekateri anketiranci raje odgovorijo z »ne vem«, kakor da preko kompleksnega procesa odgovarjanja oblikujejo ustrezen odgovor. Odgovor »ne vem« je lahko tudi odsev posameznikovega pomanjkanja motivacije za odgovarjanje in ne pomanjkanje védenja oziroma mnenja. Verjetnost za veljaven odgovor se sicer povečuje v primeru tematike vprašanj, ki jih anketiranec bolje pozna oziroma ima glede njih osebne izkušnje in se je z njimi ukvarjal v krajšem preteklem obdobju, ki je posameznikovemu spominu lažje dosegljiv. Višji deleži neodgovorov spremenljivke kot je »ne vem« se pojavljajo tudi takrat, ko anketiranec ne uspe brez težav identificirati pravilno kategorijo odgovora, ki odseva njegovo presojo ali izkušnjo. V takšnih primerih anketiranci pogosteje raje ne odgovorijo, četudi jim je na voljo npr. možnost »drugo« (Sudman in drugi 1996).

Slika 2.5: Model obdelave informacij v procesu odgovarjanja na anketna vprašanja.



Vir: Sudman in drugi 1996, 58.

Shoemaker in sodelavci (2000) v svoji raziskavi ločujejo neodgovor spremenljivke: na eni strani je odgovor »ne vem«, na drugi pa dejanske zavrnitve odgovora na zastavljeno vprašanje. Ugotavljajo, da se z večanjem občutljivosti vprašanja veča tudi delež zavrnitev odgovorov na zastavljeno vprašanje ter da se z večanjem občutljivosti vprašanja manjša delež odgovorov »ne vem«. Ugotavljajo tudi, da se z večanjem potrebnega kognitivnega napora za odgovor na zastavljeno vprašanje

veča tako delež odgovorov »ne vem« kakor tudi delež zavrnitev odgovorov na zastavljeno vprašanje (Shoemaker in drugi 2000).

3 NEODGOVOR SPREMENLJIVKE V PRIMERU ANKET O ZDRAVJU

Vpliv dejavnikov, opisanih v točki 2.2, ki vplivajo na pojav neodgovora spremenljivke, bo natančneje analiziran v primeru dveh anketnih raziskav na področju zdravja in zdravstvenega varstva. Raziskave o zdravju so namreč izredno pomemben in pogosto celo edini vir podatkov o zdravstvenem stanju in determinantah zdravja med prebivalci, zato je izrednega pomena, da kazalniki zdravja odsevajo dejanske razlike med populacijskimi skupinami. Padajoče stopnje sodelovanja, s katerimi se raziskovalci soočajo v zadnjih desetletjih, predstavljajo pomemben problem in vplivajo tudi na veljavnost in zanesljivost raziskav, povezanih z zdravjem (Ekholm 2010, 1).

Neodgovor spremenljivke lahko pomembno vpliva na izgubo informacij in če njegova obravnava ni ustrezna, lahko rezultira v znatni pristranskosti populacijskih ocen. Različne raziskave ugotavljajo, da se višji deleži neodgovorov spremenljivke v raziskavah o zdravju pojavljajo pri osebah s slabšim zdravstvenim stanjem, osebah s kroničnimi boleznimi ter osebah s kognitivnimi in telesnimi okvarami (Colsher and Wallace 1989; Guadagnoli and Cleary 1992 v Elliot in drugi 2005). Stopnje neodgovora spremenljivke v raziskavah o zdravju se višajo s starostjo anketiranca, kar kaže na povezanost starosti anketiranca s kognitivnimi ali zdravstvenimi problemi posameznega anketiranca. Na drugi strani je bila višja stopnja neodgovora enote ugotovljena pri moških, pri katerih je bila stopnja neodgovora spremenljivke na vprašanja o zdravju v primerjavi z ženskami nižja (Elliott in drugi 2005).

Beebe in sodelavci (2008) glede bodočih namer sodelovanja v anketnih raziskovanjih o zdravju pa ugotavljajo, da imajo zdravi posamezniki manjšo verjetnost, da bodo sodelovali v bodoči raziskavi o zdravju kot manj zdravi posamezniki, saj se zdijo raziskave o zdravju bolj zdravim manj pomembne, saj imajo manj za poročati. Ugotavljajo tudi večjo verjetnost sodelovanja pri krajših raziskavah.

Neodgovori spremenljivke v anketah o zdravju lahko pomembno vplivajo tudi na analizo socialno-ekonomskih neenakosti v zdravju, ki so trenutno ena najaktualnejših tem, s katerimi se soočajo raziskovalci v zdravstvu in so tudi izrednega pomena za odločevalce. »Socialno-ekonomske neenakosti v zdravju so tiste razlike v zdravstvenih stanjih med družbenimi skupinami z različnim socialno-ekonomskim statusom, ki jih lahko preprečimo in so nepravične. Za učinkovito zmanjševanje neenakosti v zdravju je nujno, da jih celovito poznamo, predvsem, kakšne so in kaj jih povzroča« (Buzeti in drugi 2011, 10). Če se v anketah o zdravju v primeru različnih socio-ekonomskih determinant, povezanih z zdravjem, pojavljajo večji deleži neodgovori spremenljivke, lahko le-ti izkrivljajo dejansko sliko. To pa lahko vodi k lažnim rezultatom ter s tem posledično tudi napačnim političnim odločitvam, saj podatki, zbrani z anketami o zdravju, predstavljajo pomemben vir tudi za politične odločevalce.

3.1 Problemi neodgovorov na področju zdravja

Vsebine vprašanj na področju zdravja se pogosto nanašajo na področja, ki jih posameznik zaznava kot občutljiva, to so npr. različna tvegana vedenja za zdravje (spolno vedenje, kajenje, uporaba alkohola in drog ipd.).

Anketna raziskovanja o zdravju in življenjskem slogu se tradicionalno uporabljajo za zbiranje podatkov o vedenjih, povezanih z zdravjem v različnih populacijskih skupinah. Ta raziskovanja so obsežnejša, pri čemer anketa traja dalj časa in običajno vključujejo naslednja vsebinska področja (McCluskey 2011):

a) kajenje,
b) uživanje alkohola,
c) indeks telesne mase,
d) zobozdravstveno stanje,
e) diete in prehranske navade,
f) uživanje drog,
g) duševno zdravje,
h) telesna dejavnost,
i) prevalenca bolezni,
j) samoporočana ocena zdravja in
k) spolna vedenja.

McCluskey (2011) ugotavlja, da v anketnih raziskovanjih o zdravju in življenjskem slogu uporaba samo enega načina anketiranja pogosto ne dosega več visokih oziroma reprezentativnih stopenj odgovora. Za zviševanje stopenj odgovorov je zato v anketnih raziskovanjih o zdravju in življenjskem slogu priporočljiva hkratna uporaba različnih načinov anketiranja.

Raziskovalci, ki zbirajo podatke za proučevanje zapletenih in občutljivih vedenj, povezanih z zdravjem, se pogosto soočajo s problemom zagotavljanja pravočasnih, veljavnih in zanesljivih podatkov, kar je še posebej povezano z vedno nižjimi stopnjami odgovorov v zadnjih desetletjih (Krosnick 1999; Singer 2006; Krosnick in drugi 2002; Krosnick 2002). Omenjena problematika je še večja, kadar gre za zbiranje podatkov o izredno mobilni populaciji, kot so npr. mladi odrasli (McMorris 2009).

Kupek (1998) v raziskavi o spolnih vedenjih ugotavlja, da se večji deleži neodgovora spremenljivke pojavljajo pri posameznikih, ki so nižje izobraženi, in tistih, ki se uvrščajo v nižji družbeni položaj, ter pri posameznikih, ki odgovarjajo v prisotnosti staršev ali drugih odraslih oseb. Na deleže neodgovorov spremenljivke vpliva tudi pomanjkanje informacij, težave s spominom oziroma priklicem dogodka ter občutljivost posameznih vprašanj.

Kaplanova in sodelavci (2001) ugotavljajo, da se višji deleži neodgovorov spremenljivke v primeru vprašanj o kadilskih navadah posameznikov pojavljajo pri podatkih, zbranih s samoanketiranjem, kot pri osebnem anketiranju. Višji deleži neodgovorov spremenljivke pri vprašanjih o kadilskih navadah so se pojavili tudi pri mlajših in manj izobraženih anketirancih ter pri daljših odsekih vprašanj, ki se nanašajo na zgodovino kadilskih navad. Tudi v splošnem velja, da so deleži neodgovorov spremenljivke na vprašanja o pitju alkoholnih pijač v primerjavi z drugimi manj občutljivimi vprašanji, višji. Glede na stopnjo izobrazbe pa se med ženskami višji deleži neodgovorov spremenljivke na vprašanja o pitju alkoholnih pijač pojavljajo pri manj izobraženih, kar kaže na to, da ženske z nižjo stopnjo izobrazbe glede občutljivih vprašanj odreagirajo drugače kot bolj izobražene ženske (Alvik 2005).

Raziskave ugotavljajo, da so podatki o pitju alkoholnih pijač v raziskavah, povezanih z zdravjem, pogosto podporočani, predvsem, kjer gre za anketiranje s pomočjo anketarja. Samoporočanje je tako trenutno najboljša metoda zbiranja podatkov o zmernem pitju alkoholnih pijač, pri čemer ne smemo pozabiti, da je samoanketiranje občutljivo na pristranskost odgovora (Alvik in drugi 2005).

Populacijske ocene o tveganih vedenjih, izračunane na podlagi podatkov, zbranih s samoanketiranjem, so zelo zanesljive, saj posamezniki lažje anonimno odgovarjajo na vprašanja o vedenjih, ki so družbeno manj sprejemljiva. Medtem ko so ocene o tveganih vedenjih, izračunane na podlagi podatkov, zbranih z osebnim anketiranjem, lahko pogosto podporočane. Samoanketiranje je tako v primerjavi z drugimi načini anketiranja (npr. z osebnim anketiranjem) učinkoviteje, kadar gre za občutljivejša vprašanja o tveganih vedenjih in različnih zdravstvenih stanjih (Kaplan 2001).

Raziskava o zdravju in zdravstvenem stanju prebivalcev na Finskem in v baltskih državah je pokazala relativno nizke stopnje neodgovora spremenljivke (od 4 % do 8 %). Višje stopnje neodgovora spremenljivke so se pojavile pri starejših in nižje izobraženih posameznikih ter pri vprašanjih o alkoholu in kadilskih navadah (Helasoja in drugi 2002).

Dengler in sodelavci (1997) v analizi neodgovora spremenljivke v raziskavi o zdravju in življenjskem slogu, v kateri so bili podatki zbrani s pomočjo poštne ankete s samoanketiranjem, ugotavljajo, da se višji deleži neodgovora spremenljivke pojavljajo proti koncu anketnega vprašalnika. Anketiranci, ki so izpustili odgovor na vsaj eno od pojasnjevalnih spremenljivk, so imeli višje deleže neodgovora enote. V splošnem so se višje stopnje neodgovora spremenljivke pojavljaje pri ženskah, medtem ko so se višje stopnje neodgovorov na specifična anketna vprašanja, npr. na vprašanja o kajenju, pojavljale pri osebah, starih od 55 do 70 let, ter

tistih z dolgotrajnimi boleznimi oziroma dolgotrajnimi zdravstvenimi težavami. Višje stopnje neodgovorov spremenljivke na vprašanje o pitju alkohola so se pojavile pri osebah, starih od 55 do 70 let, medtem ko anketiranci, ki niso odgovorili na vprašanja o zdravstvenih pregledih, pogosteje niso odgovorili tudi na vprašanje o dolgotrajnih boleznih oziroma dolgotrajnih zdravstvenih težavah ter na vprašanja o kajenju in pitju alkohola v zadnjem tednu pred anketiranjem.

Stopnje neodgovora spremenljivke pri občutljivejših tematikah se razlikujejo glede na način anketiranja. Posamezniki namreč pogosteje predčasno zaključijo z odgovarjanjem na anketo v primeru računalniško podprtega telefonskega anketiranja kot v primeru anketiranja s pomočjo anketarja (Gibble in drugi 2000).

Cohen in Duffy (2002) v raziskavi neodgovorov, opravljeni na podlagi podatkov, zbranih s poštno anketo o zdravju starostnikov, ki so jih zaradi spremljanja naknadne umrljivosti respondentov in nerespondentov povezali z britanskim centralnim registrom sistema zdravstvenega varstva, ugotavljata, da je samoocenjeno zdravstveno stanje močno povezano z bodočo umrljivostjo. In sicer je umrljivost nerespondentov višja kot umrljivost respondentov. Samoporočana ocena zdravstvenega stanja pa je povezana s poznejšo pripravljenostjo odgovarjanja.

Tudi kulturne razlike posameznikov lahko pomembno vplivajo na pojav neodgovora spremenljivke v raziskavah o zdravju. Owens in sodelavci (2001) namreč ugotavljajo obstoj kulturno pogojenih razlik med posameznimi družbenimi skupinami v stopnjah pojavnosti neodgovora spremenljivke v raziskavah o zdravju.

V raziskavah o zdravju in z zdravjem povezanih vedenjih so sicer deleži oseb, ki dejansko zavračajo odgovor na vprašanje, relativno nizke, višji pa so deleži odgovora »ne vem«, ko se anketiranci ali ne morejo natančno

spomniti določenega odgovora o svojem zdravju oziroma z zdravjem povezanimi vedenji, ali pa so odraz njihovih pomanjkljivih znanj, npr. pri vprašanjih, ki se nanašajo na zdravstveni sistem (Kauter in drugi 2006).

Najvišje stopnje neodgovora spremenljivke se tudi v anketah o zdravju pojavljajo pri vprašanju o dohodku. V ameriški »National Health Interview Survey« (NHIS) so stopnje neodgovora spremenljivke v primeru vprašanja o dohodku od leta 1998 porastle kar za 30 %, kar zmanjšuje analitično uporabnost podatkov o dohodku ter z njimi povezane zdravstvene izide in ostale podatke o zdravstvenem sistemu (Pleis in Cohen 2008).

Raziskovalci v zadnjih desetletjih ugotavljajo velik vpliv socialno-ekonomskih dejavnikov na različne izide zdravja, npr. na pričakovano trajanje življenja, umrljivost, obolevnost, manjzmožnost in doživljanje zdravja pa tudi na dostopnost in uporabo zdravstvenih storitev (Buzeti in drugi 2011).

»Sistematično zbiranje in analiziranje podatkov je predpogoj za načrtovanje in ukrepanje, za zmanjševanje neenakosti v zdravju in za spremljanje doseganja zastavljenih ciljev. Dostopnost podatkov na individualni ravni s povezavami med socialno-ekonomskimi in zdravstvenimi podatki ter razvoj novih virov podatkov sta nujna za sistematično spremljanje, analiziranje in vrednotenje neenakosti v zdravju ter vrednotenje zastavljenih ciljev« (Buzeti in drugi 2011). Za učinkovito spremljanje neenakosti v zdravju je pomemben tudi razvoj primerljivih mer družbenega položaja, ki bodo tudi mednarodno primerljive, in drugih vidikov socio-ekonomske pozicije posameznika, vezanih na individualno raven posameznika (Kreiger in drugi 1997).

Slika 3.1: Determinante zdravja.

Vir: Dahlgreen in Whitehead v Buzeti in drugi 2011.

Zgornja slika predstavlja splošne determinante zdravja. Na posameznika in njegovo zdravje vplivajo tudi bivalne in delovne razmere, dostopnost do hrane, osnovnih dobrin in storitev ter ekonomski, kulturni in okoljski vplivi širšega družbenega okolja (Buzeti in drugi 2011, 18–19). Za potrebe sistematičnega spremljanja neenakosti v zdravju so z vidika socialno-ekonomskih podatkov o prebivalcih potrebni čim bolj popolni podatki na individualni ravni, vendar se tako v administrativnih zbirkah, kakor tudi v anketnih raziskovanjih pojavljajo manjkajoči podatki na ravni posameznih spremenljivk oziroma na ravni posameznih anketnih vprašanj, ki lahko v primeru, da njihovi deleži niso zanemarljivi, izkrivljajo dejansko sliko, kar lahko vodi k lažnim rezultatom. Goyder in sodelavci (2002) ugotavljajo, da se v anketnih raziskavah med splošno populacijo, kadar se stopnja odgovora giblje nekje med 40 % in 70 % pod določenimi pogoji pojavlja »pristranskost srednjega razreda«, medtem ko se z večanjem posameznikovega socialno-ekonomskega položaja (SEP) verjetnost za odgovarjanje veča.

V letošnjem letu sta bili na področju neenakosti v zdravju v Sloveniji izdani dve znanstveni monografiji, v katerih sta bili z vidika vpliva demografskih ali socio-ekonomskih spremenljivk na neenakosti v zdravju analizirani tudi »Anketa o zdravju in zdravstvenem varstvu« in anketa »Z zdravjem povezano vedenje v šolskem obdobju«.

V raziskavi Neenakosti v zdravju in z zdravjem povezanih vedenjih slovenskih mladostnikov so bile poleg spola in starosti (11, 13 in 15 let), ki sta pomembni dimenziji socialnih razlik, upoštevane tudi neodvisne spremenljivke, ki merijo socialno-ekonomski položaj mladostnikov oziroma njihovih družin (Jeriček Klanšček in Zupanič 2011). V grobem jih lahko razdelimo v dve glavni skupini, in sicer socialno-ekonomski položaj družine ter mladostnikov osebni socialni položaj (Koivusilta in drugi 2006). Socialno-ekonomski položaj mladostnikove družine je bil merjen z lestvico »Familly Affluence Scale« (FAS), subjektivno oceno družinskega blagostanja, s tipom družine in zaposlitvenim statusom staršev. Mladostnikov osebni socialni položaj pa je bil merjen s številom res dobrih prijateljev, šolskim uspehom pri enajst- in trinajstletnikih ter programom srednješolskega izobraževanja pri petnajstletnikih (Jeriček Klanšček in Zupanič 2011).

V nadaljevanju si poglejmo deleže neodgovorov v primeru zgoraj navedenih spremenljivk, s pomočjo katerih so bile opazovane neenakosti v zdravju med mladimi. Opazimo, da delež neodgovora spremenljivke na nobeno izmed izbranih anketnih vprašanj ne presega 2 %. Najvišji delež neodgovora spremenljivke se pojavlja v primeru socialno-ekonomskega položaja družine, merjenega z lestvico FAS.

Tabela 3.1: Neodgovor spremenljivke in merjenje neenakosti v zdravju med mladimi.

Spremenljivke, s pomočjo katerih so bile opazovane neenakosti v zdravju slovenskih mladostnikov	Neodgovor spremenljivke (n)	% med vsemi odgovori
Starost.	7	0,1
Socialno-ekonomski pobožaj družine merjen z lestvico FAS.	103	1,9
Kako dobro (denarno) gre tvoji družini.	45	0,8
Tip družine.	58	1,1
Skupno število prijateljev ne glede na spol.	30	0,6
Šolski uspeh - Kako tvoj razrednik ocenjuje tvoje delo za šolo?	21	0,4
Poklic staršev.	58	1,1

Vir: HBSC 2010.

V analizi neenakosti v zdravju (Buzeti in drugi 2011), v kateri so bili uporabljeni tudi podatki, zbrani z »Anketo o zdravju in zdravstvenem varstvu«, so se raziskovalci osredotočili predvsem na razlike med skupinami oseb glede na demografske ali socio-ekonomske spremenljivke, kot so spol, starost, stopnja izobrazbe, zaposlitveni status, dohodek na družinskega člana (izračunan kot dohodek na ekvivalentnega člana gospodinjstva z uporabo prilagojene ekvivalenčne lestvice OECD) ter velikost in tip naselja. Analizirani so bili podatki o splošnem zdravstvenem stanju, uživanju sadja in zelenjave, samoporočani podatki o telesni teži in višini za izračun indeksa telesne mase, podatki o uživanju alkoholnih pijač in kajenju, o pojavljanju poškodb, ki so posledica prometnih nezgod ali nezgod v prostem času, podatki o duševnem zdravju ter izpostavljenosti težkim delovnim razmeram in posledično boleznim, ki izvirajo iz dela.

V analizah o neenakostih v zdravju (Buzeti in drugi 2011) so bili različni zdravstveni izidi opazovani glede na spol, starost, doseženo stopnjo izobrazbe, zaposlitveni status in dohodek. Ker je bil dohodek gospodinjstva izračunan kot dohodek na ekvivalentnega člana gospodinjstva z uporabo prilagojene ekvivalenčne lestvice OECD, so z vidika analize neenakosti pomembne tudi spremenljivke o številu družinskih članov in številu otrok v gospodinjstvu.

V nadaljevanju navajamo nekaj pomembnih spremenljivk s pomočjo katerih se merijo neenakosti v zdravju. »Anketa o zdravju in zdravstvenem varstvu« je podatke o spolu in starosti črpala iz vzorca, zato ti dve spremenljivki z vidika neodgovorov ne potrebujeta analize.

Ena pomembnejših spremenljivk s pomočjo katerih se merijo neenakosti v zdravju je dohodek gospodinjstva. Opažamo, da tudi v primeru združene spremenljivke o skupnem dohodku gospodinjstva, kjer sta združena odprto vprašanje in vprašanje o posameznih dohodkovnih razredih, delež neodgovorov še vedno predstavlja 25,8 % vseh odgovorov.

Tabela 3.2: Neodgovor spremenljivke in merjenje neenakosti v zdravju med odraslo populacijo.

Spremenljivki s pomočjo katerih so bile opazovane neenakosti v zdravju med odraslo populacijo[1]	Neodgovor spremenljivke (n)	% med vsemi odgovori
Združena spremenljivka: Kakšen je skupni mesečni neto dohodek/prejemek vašega gospodinjstva?	546	25,8
Število otrok v gospodinjstvu, starih štiri leta ali manj.	18	0,8

Vir: EHIS 2007.

V nadaljevanju bodo predstavljeni še neodgovori, vezani na gospodinjstvo, in sicer v primeru vprašanja o skupnem številu oseb v gospodinjstvu, številu otrok v gospodinjstvu glede na starost otrok in številu drugih oseb v gospodinjstvu glede na njihovo starost ter tip gospodinjstva. Zgoraj navedena spremenljivka o številu otrok v gospodinjstvu, ki predstavlja eno izmed spremenljivk o številu družinskih članov, so bile izračunane na podlagi relativno kompleksne matrike gospodinjstev. Na vprašanje o skupnem številu oseb v gospodinjstvu so odgovorili prav vsi anketiranci.

[1] Navedeni sta le, v analizi uporabljeni spremenljivki, kjer se je pojavil neodgovor spremenljivke.

Iz matrik gospodinjstev je bilo mogoče izračunati vse zahtevane spremenljivke za potrebe Eurostata. Neodgovor se je namreč pojavil samo v primeru 18 anketirancev, kar predstavlja 0,8 % vseh odgovorov. Zgoraj je predstavljena le ena izmed izračunanih spremenljivk o številu otrok v gospodinjstvu glede na njihovo starost[2].

Tudi spremenljivko tip gospodinjstva je bilo mogoče izračunati za vse anketirance. Glede na to, da so se na podlagi matrike gospodinjstev izračunanih spremenljivk neodgovori pojavili samo pri spremenljivkah o skupnem številu otrok glede na starost otrok v gospodinjstvu in številu drugih oseb glede na njihovo starost, bodo v nadaljevanju z vidika neodgovora spremenljivke obravnavane samo spremenljivke, vezane na gospodinjstvo, v katerem posameznik živi.

Vzorci manjkajočih odgovorov v »Anketi o zdravju in zdravstvenem varstvu« so celostno obravnavani v poglavju 4.3.

[2] Delež neodgovora je enak tudi v primeru ostalih spremenljivk o številu otrok v gospodinjstvu, glede na njihovo starost ter številu drugih članov gospodinjstva.

4 EMPIRIČNA ANALIZA NEODGOVORA SPREMENLJIVKE

V pričujočem poglavju bosta najprej opisani obe raziskavi o zdravju na podlagi katerih bodo izvedene analize neodgovorov na anketno vprašanje, nato pa bomo obravnavali neodgovor spremenljivke med otroki in mladimi, socio-demografske značilnosti nerespondentov, tehnike reševanja problema neodgovora spremenljivke in vpliv načina anketiranja na stopnje neodgovora spremenljivke.

4.1 Opis izbranih raziskav

V nadaljevanju besedila bo sledil opis obeh anketnih raziskav o zdravju, v primeru katerih bosta tudi empirično preverjeni raziskovalni tezi.

Anketa o zdravju in zdravstvenem varstvu

»Anketa o zdravju in zdravstvenem varstvu« (angl. »European Health Interview Survey«, v nadaljevanju EHIS) je bila v Sloveniji prvič izvedena konec leta 2007, predvidoma pa naj bi se jo izvajalo vsakih 5 let. V metodologiji (načinu vzorčenja, vsebini vprašalnikov) slovenske »Ankete o zdravju in zdravstvenem varstvu« 2007 so bila upoštevana priporočila Eurostata. EHIS je v letih 2007 do 2010 izvedla večina držav Evropske unije, zato so njeni izsledki tudi mednarodno primerljivi. Enoto opazovanja so predstavljali prebivalci Slovenije, stari 15 let ali več, ki živijo v zasebnih gospodinjstvih (niso institucionalizirani). V vzorec je bilo vključenih 3400 oseb, ki so bile na dan začetka anketiranja (20. oktobra 2007) stare 15 let ali več. Osnovo za vzorčni okvir je predstavljal okvir popisnih okolišev in Centralni register prebivalstva. Uporabljeno je bilo dvostopenjsko vzorčenje, in sicer je bil uporabljen stratificiran dvostopenjski vzorec (PPS s ponavljanjem), kjer je bilo eksplicitno stratificirano po velikosti in tipu naselja, implicitno pa po statističnih

regijah. Na prvi stopnji je bilo izbranih 425 vzorčnih enot (skupin popisnih okolišev), nato pa v vsaki izbrani vzorčni enoti po 8 oseb, starih 15 ali več. Končna velikost vzorca je bila 3387 oseb. Večina podatkov je bila zbranih z osebnim anketiranjem na naslovih izbranih oseb z vprašalnikom na papirju (PAPI), podatki o kajenju, alkoholu in drogah ter podatki o samoplačniških stroških pa so bili zbrani s samoanketiranjem z vprašalnikom na papirju. Stopnja odgovora v anketi je bila 68 % (IVZ 2008).

Anketa »Z zdravjem povezano vedenje v šolskem obdobju«

Anketa »Z zdravjem povezano vedenje v šolskem obdobju« (angl. »Health Behaviour Among School Children«, v nadaljevanju HBSC) je bila leta 2010 v Sloveniji izvedena že tretjič, vanjo pa je bilo vključenih 43 držav. V raziskavo so bili zajeti le všolani mladostniki, ne pa tudi t. i. osipniki, zato je pri interpretaciji in posploševanju dobljenih rezultatov potrebna previdnost. Anketa je bila izvedena s standardiziranim mednarodnim vprašalnikom na reprezentativnem vzorcu slovenskih všolanih enajst-, trinajst- in petnajstletnikov, anketiranje pa je bilo izvedeno ob pomoči šolskih svetovalnih delavcev ali učiteljev v februarju in marcu 2010. Vprašalnike na papirju so učenci izpolnjevali samostojno, izpolnjene vprašalnike pa so nato šolski svetovalni delavci ali učitelji po pošti posredovali na Inštitut za varovanje zdravja RS, kjer so bili pregledani in vnešeni v bazo. Pregledana in po mednarodni metodologiji pripravljena baza podatkov je bila nato preko spletne aplikacije posredovana v mednarodni center na Norveškem, kjer so zbrane podatke prečistili skladno z mednarodno določenimi pravili (Jeriček Klanšček in Zupanič 2011).

4.2 Posebnosti anketiranja otrok in mladostnikov

Proces spraševanja in odgovarjanja je ključen za uspešno izvedbo anketne raziskave. Kadar anketiramo otroke in mladostnike, se moramo zavedati kognitivnega razvoja otroka, ki ima velik vpliv na proces spraševanja in odgovarjanja. Stopnja razvoja otroka ima namreč vpliv na njihovo izvajanje nalog, potrebnih za kreiranje odgovora na zastavljeno vprašanje. Otroci morajo namreč najprej razumeti vprašanje in ugotoviti njegov pomen, nato v svojem spominu priklicati potrebne informacije in jih uporabiti v pripravi okvirnega odgovora, nazadnje pa izbrati ustrezno ponujeno možnost odgovora. V celotnem procesu odgovarjanja poteka tudi evalvacija možnih odgovorov v smislu družbene zaželenosti in situacijske primernosti. Otroci se namreč vedejo različno ob prisotnosti svojih učiteljev, staršev in sovrstnikov (Borgers in drugi 2000).

Otroke in mladostnike, stare od 7 do 18 let, je že primerno raziskovati s posebnimi anketnimi raziskovanji. Medtem ko otroci mlajši od 7 let še nimajo razvitih ustreznih kognitivnih spretnosti, da bi jih lahko učinkovito in sistematično anketirali, pa se v starosti sedmih let pri otroku zgodi velik napredek v kognitivnem in družbenem dozorevanju, zaradi česar so otroci pri tej starosti že sposobni samoizpolnjevati anketne vprašalnike na papirju (De Leeuw 2011). Raziskave so pokazale, da sposobnost branja pri otrocih pomembno vpliva na stopnje neodgovora spremenljivke ter kakovost zbranih podatkov. Otroci z bolj razvitimi govornimi in bralnimi spretnostmi bolje razumejo vprašanja, zaradi česar je kakovost njihovih odgovorov višja (Borgers in drugi 2000). Tudi občutljivost anketnih vprašanj lahko pri otrocih pomembno vpliva na deleže neodgovora spremenljivke, in sicer se nižje stopnje neodgovorov na občutljiva anketna vprašanja pojavljajo pri mlajših otrocih, medtem ko so deleži manjkajočih odgovorov na občutljiva vprašanja pri mladostnikih višji (Borgers in Hox 2001).

Raziskave ugotavljajo, da na neodgovor spremenljivke pomembneje vplivajo značilnosti otrok in mladostnikov, kot pa značilnosti anketnih vprašanj. Višji deleži neodgovora spremenljivke se pojavljajo med mlajšimi otroki. Med značilnostmi anketnih vprašanj v raziskovanjih med otroki in mladostniki pa ima na neodgovor spremenljivke pomemben vpliv število ponujenih odgovorov na zastavljeno vprašanje. Namreč večje kot je število možnih odgovorov na posamezno anketno vprašanje, večje so stopnje neodgovorov. Tudi nejasnost zastavljenih vprašanj lahko pomembno vpliva na neodgovore, in sicer so stopnje odgovorov višje v primeru vprašanj, pri katerih so merske lestvice dvoumne. Slednje se sicer izraziteje kaže pri starejših otrocih in mladostnikih (Borgers in Hox 2001).

Epidemiološke raziskave o kajenju mladostnikov se večinoma zanašajo na samoporočane prevalenčne stopnje in vzorce kajenja (Griesler in drugi 2008). Na te samoporočane prevalenčne stopnje kajenja, uporabe prepovedanih drog in drugih občutljivih, tveganih in nelegalnih vedenj pa vpliva tudi kraj anketiranja (šola v primerjavi z gospodinjstvom) in način anketiranja (samoanketiranje v primerjavi z anketiranjem s pomočjo anketarja in anketiranje na papirju proti računalniško podprtemu anketiranju) (Brener in drugi 2006; Frendrich in Johnson 2001; Tourangeau in Smith 1996). Vpliv kraja in načina anketiranja izvira iz posameznikovih zaznavanj zaupnosti in zasebnosti ter družbeno zaželenih odgovorov (Griesler in drugi 2008).

Tudi umeščenost anketnih vprašanj v vprašalniku in njihova tematika pomembno vplivata na stopnje neodgovorov v raziskavah o tveganih vedenjih med mladimi. Analize namreč kažejo, da imajo vprašanja, ki so predhodna vprašanjem o spolnih vedenjih, in vprašanja, ki tem sledijo, nižje stopnje neodgovora spremenljivke, kot vprašanja o spolnih vedenjih. Medtem ko imajo anketna vprašanja s podobno tematiko kot vprašanja o

spolnih vedenjih tudi podobne stopnje neodgovora spremenljivke (Saewyc in drugi 2004).

Empirična analiza HBSC

V nadaljevanju prikazana analiza manjkajočih vrednosti podatkov je bila pripravljena s programom SPSS 19, in sicer s posebnim modulom za analizo manjkajočih vrednosti, ki omogoča tudi grafični prikaz vzorcev neodgovorov. Zaradi preglednosti prikazov v nadaljevanju besedila smo nabor izbranih spremenljivk skrčili smiselno, glede na posamezne vsebinske sklope in glede na vidik občutljivosti posameznih anketnih vprašanj. Podatki so analizirani samo na vzorcu 15-letnikov, saj 11- in 13-letniki niso odgovarjali na nekatera vprašanja o tveganih vedenjih (nekatera analizirana vprašanja o pitju alkoholnih pijač, kajenju marihuane, spolnih odnosih). V spodnji tabeli so predstavljeni deleži neodgovorov spremenljivke na izbrana anketna vprašanja. Kot lahko opazimo se največja deleža neodgovorov pojavita pri vprašanjih o spolnih odnosih, sledita vprašanji o socialno-ekonomskem položaju družine in rekreaciji v prostem času.

Tabela 4.1: Neodgovor spremenljivke na posamezna izbrana anketna vprašanja v HBSC – 1. del.

Izbrana anketna vprašanja zbrana s samoanketiranjem	n	Neodgovor spremenljivke	
		n	%
Koliko si bil/a star/a, ko si prvič imel/a spolne odnose? (m46)	1757	58	3,2
Ali že kdaj imel/a spolne odnose ...? (m45)	1762	53	2,9
Ali si že kdaj kadil/a marihuano (travo)?... V vsem življenju? (m42)	1768	47	2,6
Socialno ekonomski položaj družine (fas_novo)	1788	27	1,5
Kako pogosto se običajno rekreiraš v prostem času (izven pouka) tako intenzivno, da ti zmanjka sape ali se prepotiš? (m17)	1795	20	1,1
Zaposlitveni status staršev (SES_st)	1796	19	1
Tip družine (druzina_m)	1799	16	0,9

Vir: HBSC 2010.

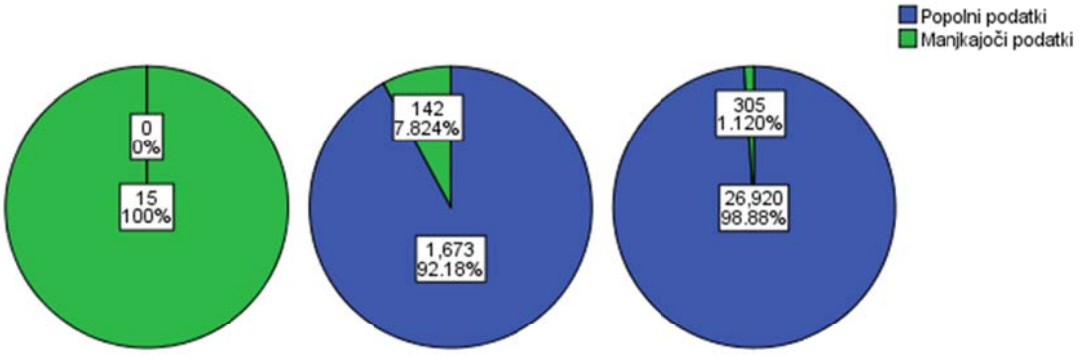
Tabela 4.2: Neodgovor spremenljivke na posamezna izbrana anketna vprašanja v HBSC – 2. del.

Izbrana anketna vprašanja zbrana s samoanketiranjem	n	Neodgovor spremenljivke	
Kako pogosto piješ vino? (m28)	1803	12	0,7
Kako pogosto piješ žgane pijače? (m29)	1802	13	0,7
Kako pogosto piješ mešane gazirane alkoholne pijače? (m30)	1802	13	0,7
Ali si že spil/a toliko alkoholne pijače, da si bil/a prav zares pijan/a? (m35)	1806	9	0,5
Kako pogosto piješ pivo? (m27)	1809	6	0,3
Tedensko pitje alkohola (alkoweek)	1810	5	0,3
Kako pogosto ješ sadje? (m7)	1811	4	0,2
Kako pogosto si umivaš zobe? (m15)	1812	3	0,2

Spodnja slika prikazuje splošen pregled neodgovora spremenljivke. Opazimo lahko naslednje:

- vsaj en neodgovor se pojavlja pri vseh izbranih anketnih vprašanjih;
- vsaj en neodgovor spremenljivke ima 7,8 % izmed vseh 1815 respondentov;
- 305 vrednosti (respondenti x anketna vprašanja) izmed vseh 2997 vrednosti, je manjkajočih.

Slika 4.1: Splošen pregled neodgovora spremenljivke pri izbranih anketnih vprašanjih v HBSC.

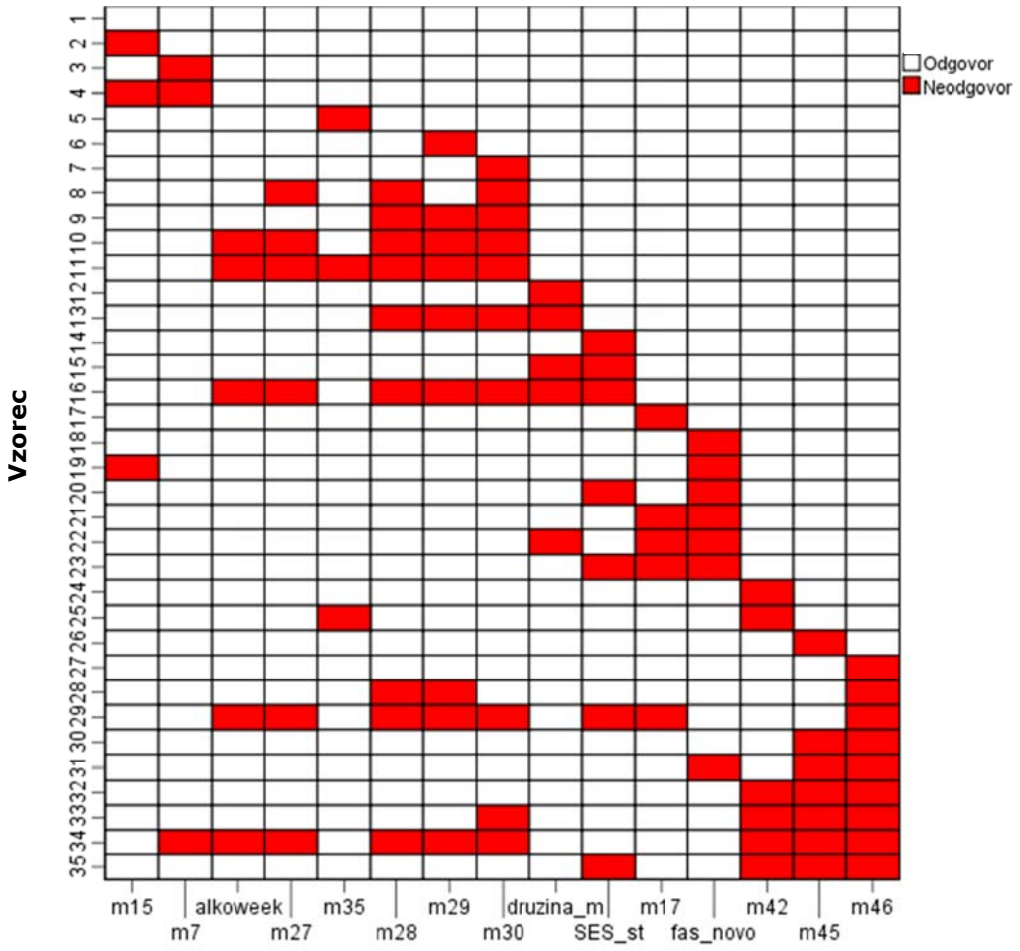


Slika na naslednji strani prikazuje vzorec neodgovora spremenljivke, glede na izbrana anketna vprašanja. Vsak vzorec neodgovora

spremenljivke ustreza skupini respondentov s popolnimi podatki (odgovori) in/ ali neodgovori npr.:

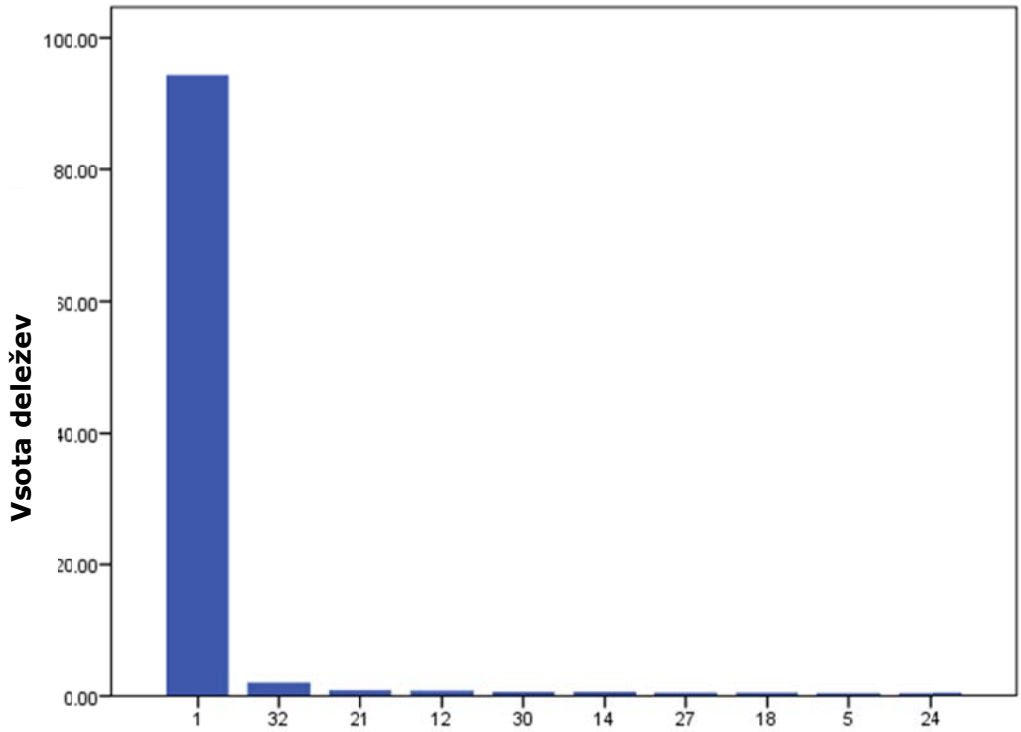
- Vzorec 2 predstavlja respondente, ki so niso odgovorili na vprašanje o pogostosti umivanja zob;
- Vzorec 9 respondente, ki niso odgovorili na vprašanja o tem kako pogosto pijejo pivo, vino in žgane pijače;
- Vzorec 11 respondente, ki niso odgovorili na nobeno izmed vprašanj o pitju alkoholnih pijač;
- Vzorec 23 respondente, ki niso odgovorili na vprašanja poklicu staršev, rekreaciji v prostem času in socialno-ekonomskem položaju družine;
- Vzorec 29 predstavlja respondente, ki niso odgovorili na vsa vprašanja o pitju alkoholnih pijač, vprašanje o poklicu staršev, rekreaciji ter vprašanje o starosti ob prvem spolnem odnosu;
- Vzorec 35 predstavlja respondente, ki niso odgovorili na vprašanje o poklicu staršev, uživanju marihuane ter na obe vprašanji o spolnih odnosih;
- Vzorec 1 pa respondente, ki so odgovorili na vsa izbrana anketna vprašanja.

Slika 4.2: Vzorec neodgovorov spremenljivke pri izbranih vprašanjih HBSC.



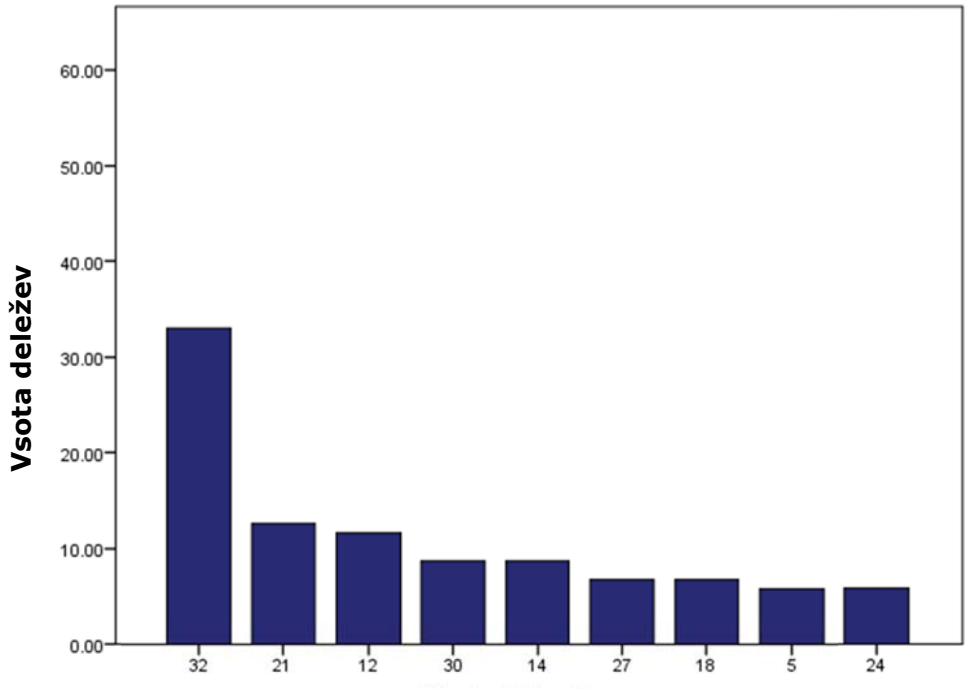
Poglejmo si še 10 najpogosteje pojavljajočih se vzorcev v izbranem setu podatkov. Opazimo lahko, da več kot 95 % vseh respondentov pripada vzorcu 1 (respondenti, ki so odgovorili na vsa anketna vprašanja). Ostalim vzorcem pa pripada manj kot 1 % vseh respondentov.

Slika 4.3: Najpogosteje pojavljajoči vzorci neodgovora spremenljivke v podatkih HBSC.



V primeru, ko opazujemo vzorce neodgovorov spremenljivke ločeno od vzorca 1, kjer so vsi respondenti odgovorili na vsa anketna vprašanja, lahko opazimo, da dobrih 30 % vseh respondentov pripada vzorcu pripada vzorcu 32, ki predstavlja posameznike, ki niso odgovorili na dve vprašanji o spolnih odnosih in vprašanje o kajenju marihuane, 12 % vseh respondentov pripada vzorcu 21 in predstavlja posameznike, ki niso odgovorili na vprašanji o rekreaciji in socialno-ekonomskem položaju družine, 11 % vseh respondentov pa vzorcu 12 in predstavlja posameznike, ki niso odgovorili na vprašanje o tipu družine. Ostalim vzorcem pa pripada manj kot 10 % vseh respondentov.

Slika 4.4: Najpogosteje pojavljajoči vzorci neodgovora spremenljivke v podatkih HBSC, brez najpogosteje ponavljajočega vzorca.



Zgornje ugotovitve dobljene z analizo neodgovora spremenljivke po posameznih izbranih anketnih vprašanjih so nas presenetile, saj smo predvidevali, da bodo stopnje neodgovora spremenljivke v primeru podatkov zbranih s samoanketiranjem z vprašalnikom na papirju med mladimi, bistveno višje.

4.3 Neodgovor spremenljivke med odraslo populacijo

Neodgovori spremenljivke med odraslo populacijo so obravnavani že v drugem in tretjem poglavju, zato se bomo v nadaljevanju pričujočega poglavja osredotočili zgolj na analizo neodgovorov izbranih vprašanj iz »Ankete o zdravju in zdravstvenem varstvu«. Z vidika neodgovora spremenljivke bodo obravnavana naslednja anketna vprašanja: vprašanja o splošnem zdravstvenem stanju ter ostala ključna vprašanja v posameznih sklopih vprašalnika ter vprašanja o alkoholu in dohodku. Podatki o zdravstvenem stanju, determinantah zdravja in sistemu zdravstvenega varstva so bili zbrani s pomočjo anketarja z vprašalnikom na papirju, podatki o alkoholu pa so bili zbrani s samoanketiranjem z vprašalnikom na papirju. Vprašanja o alkoholu so namreč smatrana kot izredno občutljiva, deleži poročanega pitja alkoholnih pijač pa pri zbiranju podatkov s pomočjo anketarja tudi v veliki meri podcenjeni, zato je bilo treba pri izvedbi ankete upoštevati priporočilo Eurostata, da se pri teh vprašanjih uporabi zbiranje podatkov s samoanketiranjem z vprašalnikom na papirju. Anketarji so imeli navodilo, da vprašalnik na papirju s sklopoma vprašanj o »Samoplačniških stroških« in »Kajenju, alkoholu in drogah«, anketiranci izpolnijo sami. Anketar je pred pričetkom anketiranja v desni zgornji kot ankete vpisal šifro izbrane osebe, ter poleg vprašalnika za samoanketiranje anketirancu izročil tudi kuverto, v katero je nato anketiranec vložil izpolnjen vprašalnik (Lavtar in drugi 2007).

Kveder (2005) je opredelil različne ravni posameznih stopenj odgovora, in sicer *kazalnik neodgovora anketnega vprašanja*, kjer je izid posameznega anketnega vprašalnika dihotomna spremenljivka, kjer odgovoru pripišemo oznako 1, neodgovoru pa oznako 0:

$$\rho = \begin{bmatrix} 1 & \text{pomeni veljaven odgovor} \\ 0 & \text{pomeni neodgovor} \end{bmatrix}.$$

Ravni posameznih stopenj odgovora lahko izračunamo na treh ravneh opazovanja, in sicer na ravni anketiranca, anketarja in na ravni anketnega vprašanja, pri čemer je struktura podatkov naslednja (Kveder 2005):

- $R_i, i = 1..n^R$ pomeni anketirance,
- $I_j, j = 1..n^I$ pomeni anketarje,
- $Q_k, k = 1..n^Q$ pomeni anketna vprašanja,
- n^R pomeni število anketirancev,
- n^J pomeni število anketarjev,
- n^Q pomeni skupno število različnih vprašanj,
- $\rho_{i,j,k}$ pomeni izid anketiranja za posameznega anketiranca R_i na vprašanje Q_k, ki ga je postavil anketar I_j,
- n_i^Q pomeni skupno število vprašanj, ki so bila zastavljena anketirancu R_i,
- n_j^R pomeni skupno število anketirancev dodeljenih posameznemu anketarju I_j,
- n_k^I pa pomeni skupno število anketarjev, ki so vprašali vprašanje Q_k.

Stopnjo odgovora na ravni posameznega anketiranca izračunamo na naslednji način (Kveder 2005):

$$RR_i^R = \frac{\sum_{k=1}^{n_i^Q} \rho_{i,k}}{n_i^Q},$$

kjer n_i^Q pomeni število vseh opazovanih vprašanj, na katera bi anketiranec moral odgovoriti. Zgornji izraz se sicer posploši, saj hierarhično izključevanje anketirancev pomeni, da vsakemu anketirancu postavlja vprašanja samo en anketar.

Stopnjo odgovora na ravni posameznega anketarja izračunamo na naslednji način (Kveder 2005):

$$RR_j^I = \frac{\sum_{i=1}^{n_j^R} \sum_{k=1}^{n_i^Q} \rho_{i,j,k}}{\sum_{i=1}^{n_j^R} n_{i,j}^Q},$$

stopnjo odgovora na ravni posameznega anketnega vprašaja pa kot:

$$RR_k^Q = \frac{\sum_{j=1}^{n_k^I} \sum_{i=1}^{n_j^R} \rho_{i,j,k}}{\sum_{j=1}^{n_k^I} n_{j,k}^R}.$$

Skupna stopnja odgovorov predstavlja vsoto vseh veljavnih odgovorov glede na vse anketirance, anketarje in vsa opazovana vprašanja (Kveder 2005):

$$RR = \frac{\sum_{j=1}^{n^I} \sum_{i=1}^{n_j^R} \sum_{k=1}^{n_{i,j}^Q} \rho_{i,j,k}}{\sum_{j=1}^{n^I} \sum_{i=1}^{n_j^R} n_{i,j}^Q}.$$

V nadaljevanju bodo na primeru »Ankete o zdravju in zdravstvenem varstvu« najprej predstavljeni deleži neodgovora spremenljivke po posameznih izbranih anketnih vprašanjih, nato pa izračunane posamezne, zgoraj obravnavane stopnje odgovora na ravni posameznega anketnega vprašanja.

Empirična analiza neodgovora spremenljivke na primeru »Ankete o zdravju in zdravstvenem varstvu«

V pričujočem poglavju so prestavljene neutežene frekvenčne porazdelitve izbranih vprašanj, in sicer vprašanj o splošnem zdravstvenem stanju, ostalih ključnih vprašanj v posameznih sklopih vprašalnika ter vprašanj o alkoholu in dohodku. Podatki o zdravstvenem stanju, determinantah zdravja in sistemu zdravstvenega varstva so bili zbrani s pomočjo anketarja z vprašalnikom na papirju, podatki o alkoholu pa so bili zbrani s samoanketiranjem z vprašalnikom na papirju. Podatki o alkoholu in uživanju marihuane so namreč opredeljeni kot izredno občutljivi, deleži poročanega pitja alkoholnih pijač pa v primeru zbiranja podatkov s pomočjo anketarja tudi v veliki meri podcenjeni, zato je bilo pri izvedbi ankete treba upoštevati priporočilo Eurostata, da se pri teh vprašanjih uporabi zbiranje podatkov s samoanketiranjem.

Spodnja tabela prikazuje neodgovor spremenljivke v primeru vprašanj o splošnem zdravstvenem stanju.

Tabela 4.3: Neodgovori spremenljivke v primeru vprašanj o splošnem zdravstvenem stanju v EHIS 2007.

Izbrana anketna vprašanja zbrana s pomočjo anketarja z vprašalnikom na papirju	Neodgovor spremenljivke (n)	Neodgovor spremenljivke na ravni posameznega anketnega vprašanja (%)
Kakšno je vaše splošno zdravstveno stanje?	2	0,1
Ali imate kakšno dolgotrajno bolezen ali dolgotrajno zdravstveno težavo?	5	0,2
V kolikšni meri ste zadnjih 6 mesecih ali dlje ovirani zaradi zdravstvenih težav pri običajnih aktivnostih?	6	0,3

Vir: EHIS 2007.

Skupen delež manjkajočih odgovorov se v sklopu splošnega zdravstvenega stanja pojavlja v primeru dolgotrajne oviranosti zaradi zdravstvenih težav pri običajnih aktivnostih, kjer znaša 0,3 %.

V spodnji tabeli so navedeni podatki za nekaj *izbranih* bolezni in bolezenskih stanj iz seznama bolezni in bolezenskih stanj vprašalnika.

Tabela 4.4: Neodgovori spremenljivke v primeru vprašanj o izbranih boleznih in bolezenskih stanjih v EHIS 2007.

Izbrana anketna vprašanja zbrana s pomočjo anketarja z vprašalnikom na papirju	Neodgovor spremenljivke (n)	Neodgovor spremenljivke na ravni posameznega anketnega vprašanja (%)
Ali imate oz. ste že kdaj imeli katero izmed naslednjih bolezenskih stanj? Astma.	2	0,1
Ali imate oz. ste že kdaj imeli katero izmed naslednjih bolezenskih stanj? Kronični bronhitis.	3	0,1
Ali imate oz. ste že kdaj imeli katero izmed naslednjih bolezenskih stanj? Srčni infarkt.	2	0,1
Ali imate oz. ste že kdaj imeli katero izmed naslednjih bolezenskih stanj? Bolečine v križu.	2	0,1
Ali imate oz. ste že kdaj imeli katero izmed naslednjih bolezenskih stanj? Sladkorna bolezen.	3	0,1
Ali imate oz. ste že kdaj imeli katero izmed naslednjih bolezenskih stanj? Rak (maligni tumor, vključena tudi levkemija in limfom).	2	0,1
Ali imate oz. ste že kdaj imeli katero izmed naslednjih bolezenskih stanj? Druga dolgotrajna bolezen.	8	0,4

Vir: EHIS 2007.

Skupni delež neodgovora spremenljivke je najvišji v primeru »druge dolgotrajne bolezni«, kjer je manjkajočih 8 vrednosti, kar predstavlja 0,4 % vseh vrednosti. Pri vprašanju o jemanju zdravil v zadnjih dveh tednih se neodgovor spremenljivke ni pojavil.

Spodaj so prikazani neodgovori spremenljivke v primeru vprašanj o pitju alkoholnih pijač in uživanju marihuane. Podatki so bili zbrani *s samoanketiranjem z vprašalnikom na papirju*. Opazimo lahko, da osnovno vprašanje o pitju alkoholnih pijač anketirancem očitno ni zbujalo pretiranega nelagodja, saj beležimo le 0,2 % neodgovorov na to vprašanje, medtem ko so neodgovori v primeru pitja alkoholnih pijač ob posameznih dnevih tedna med tistimi, ki pijejo alkoholne pijače vsaj 2- do 3-krat na teden ali pogosteje, bistveno višji in segajo celo do 39,4 %. Anketiranci očitno kot najmanj problematično zaznavajo pitje alkoholnih pijač ob petkih in sobotah, kjer so deleži neodgovora najnižji med posameznimi dnevi v tednu (okoli 20 %).

V primeru vprašanja o popitih šestih mericah alkoholnih pijač v zadnjih 12 mesecih med tistimi, ki so v zadnjih 12 mesecih alkoholne pijače pili vsaj nekajkrat letno neodgovori spremenljivke porastejo celo na 46,2 %, kar gre pripisati tako tematiki vprašanja, načinu anketiranja ter preskokom vprašanja (anketiranci, ki niso vešči uporabe preskokov v primeru posameznih anketnih vprašanj, lahko tudi po pomoti določena anketna vprašanja izpustijo). Na drugi strani je neodgovor spremenljivke na vprašanje o kajenju marihuane nizek in predstavlja samo 0,3 % vseh odgovorov.

Tabela 4.5: Neodgovori spremenljivke v primeru vprašanj pitju alkoholnih pijač in uživanju marihuane v EHIS 2007.

Izbrana anketna vprašanja zbrana s samoanketiranjem z vprašalnikom na papirju	Neodgovor spremenljivke (n)	Neodgovor spremenljivke na ravni posameznega anketnega vprašanja (%)
Kako pogosto ste v zadnjih 12 mesecih pili kakršnekoli alkoholne pijače?	5	0,2
Koliko pijač, ki vsebujejo alkohol ste popili vsak dan v tipičnem tednu, kadar ste pili alkoholne pijače? Ponedeljek.	140	32,9
Koliko pijač, ki vsebujejo alkohol ste popili vsak dan v tipičnem tednu, kadar ste pili alkoholne pijače? Torek.	165	38,7
Koliko pijač, ki vsebujejo alkohol ste popili vsak dan v tipičnem tednu, kadar ste pili alkoholne pijače? Sreda.	146	34,3
Koliko pijač, ki vsebujejo alkohol ste popili vsak dan v tipičnem tednu, kadar ste pili alkoholne pijače? Četrtek	171	40,1
Koliko pijač, ki vsebujejo alkohol ste popili vsak dan v tipičnem tednu, kadar ste pili alkoholne pijače? Petek.	91	21,4
Koliko pijač, ki vsebujejo alkohol ste popili vsak dan v tipičnem tednu, kadar ste pili alkoholne pijače? Sobota.	88	20,7
Koliko pijač, ki vsebujejo alkohol ste popili vsak dan v tipičnem tednu, kadar ste pili alkoholne pijače? Nedelja.	132	31,0
Kako pogosto ste v zadnjih 12 mesecih popili 6 ali več meric pijače ob eni priložnosti?	755	46,2
Ali ste v zadnjih 12 mesecih uživali kanabis?	6	0,3

Vir: EHIS 2007.

V nadaljevanju bodo predstavljena še vprašanja o dohodku, in sicer bodo najprej predstavljena vprašanja in deleži neodgovora spremenljivke na vprašanje o vrstah dohodkov in prejemkov gospodinjstva, nato pa še obe vprašanji o skupnem dohodku gospodinjstva (odprto vprašanje, kjer so morali anketiranci navesti natančen znesek in vprašanje, kjer so se anketiranci umestili v določen dohodkovni razred – to vprašanje je bilo postavljeno samo tistim anketirancem, ki na odprto vprašanje o dohodku gospodinjstva ali niso znali odgovoriti, ali pa so zavračali odgovor na to vprašanje).

Tabela 4.6: Neodgovori spremenljivke v primeru vprašanj o vrstah dohodkov in skupnem dohodku gospodinjstva v EHIS 2007.

Izbrana anketna vprašanja zbrana s pomočjo anketarja z vprašalnikom na papirju	Neodgovor spremenljivke (n)	Neodgovor spremenljivke na ravni posameznega anketnega vprašanja (%)
Dohodek od zaposlitve (kot zaposleni ali samozaposleni).	13	0,6
Podpora za brezposelne.	14	0,7
Starostna ali družinska pokojnina.	14	0,7
Dodatki za bolezen in invalidnost.	14	0,7
Družinski/ otroški dodatki.	21	1,0
Stanovanjski dodatki.	17	0,8
Štipendija.	14	0,7
Drugi redni prejemki.	17	0,8
Brez vira dohodka.	17	0,8
Ali veste, kakšen je skupni mesečni neto dohodek/prejemek vašega gospodinjstva?	81	3,8
Kakšen je skupni mesečni neto dohodek/prejemek vašega gospodinjstva?[3]	272	26,7
Ali lahko namesto natančnega zneska ocenite vaše skupne neto dohodke/prejemke in jih razvrstite v enega izmed spodaj navedenih razredov?	438	34,7

Vir: EHIS 2007.

V primeru dohodkov oziroma prejemkov, ki se nanašajo bolj na posameznika, kot so npr. dohodek od zaposlitve, podpora za brezposelne, starostna pokojnina ipd., neodgovori predstavljajo tiste, ki zavračajo odgovor na zastavljeno vprašanje, medtem ko se v primeru dohodkov oziroma prejemkov, ki se nanašajo bolj na druge člane gospodinjstva, kot so npr. otroški dodatek in stanovanjski dodatki, med manjkajočimi pojavijo tudi odgovori »ne vem«. V splošnem delež neodgovorov na vprašanja o posameznih dohodkih oziroma prejemkih gospodinjstva ne presega 1 %. Neodgovor spremenljivke, na drugi strani pa v primeru vprašanja o poznavanju skupnega neto dohodka gospodinjstva, že

[3] Odprto vprašanje.

predstavlja 4 % vseh odgovorov. Neodgovor spremenljivke se v primeru vprašanj o dohodku gospodinjstva še povečuje. Kar 26,7 % anketirancev tistih, ki poznajo točen dohodek gospodinjstva, tega ni želelo navesti. Podobno tudi 34,7 % tistih, ki so zavračali odgovor na vprašanje o dohodku, ali pa niso poznali odgovora ni odgovorilo na vprašanje o umestitvi v posamezne dohodkovne razrede.

Opazimo lahko torej, da se največji delež manjkajočih odgovorov pojavlja v primeru vprašanja tveganih oziroma družbeno nezaželenih vedenj, torej v primeru popitih šestih mericah alkoholnih pijač ob eni priložnosti v zadnjih 12 mesecih, kjer neodgovori predstavljajo skoraj polovico vseh odgovorov, kjer so bili odgovori zbrani s samoanketiranjem z vprašalnikom na papirju. Visoki deleži neodgovorov so tudi v primerih pitja alkoholnih pijač v posameznih dnevih tedna (do 39,4 %). Medtem ko vprašanja o posameznih dohodkih gospodinjstva z vidika neodgovorov niso problematična, saj neodgovori v nobenem primeru ne presegajo 1 %, pa vprašanja o skupnem dohodku gospodinjstva očitno pri anketirancih zbujajo nepripravljenost odgovarjanja. Deleži neodgovorov so namreč visoki, saj tudi pri združeni spremenljivki, kjer sta združena odprto vprašanje in vprašanje o posameznih dohodkovnih razredih, delež neodgovora spremenljivke še vedno predstavlja četrtino vseh odgovorov. Neodgovori spremenljivke pri ostalih vprašanjih, npr. o splošnem zdravstvenem stanju, posameznih izbranih boleznih in bolezenskih stanjih, zdravilih ipd., ne presegajo 1 %.

V nadaljevanju prikazana analiza neodgovora spremenljivke je bila izvedena s programom SPSS 19, in sicer s posebnim modulom za analizo manjkajočih vrednosti, ki omogoča tudi grafični prikaz vzorcev neodgovorov. Zaradi preglednosti prikazov v nadaljevanju besedila smo nabor izbranih spremenljivk skrčili smiselno glede na posamezne vsebinske sklope in glede na način na katerega so bili podatki zbrani.

V tabeli na naslednji strani so predstavljeni deleži neodgovorov spremenljivke na *izbrana anketna vprašanja*. Kot lahko opazimo se največji delež neodgovora spremenljivke pojavi pri vprašanju o popitih vsaj šestih mericah alkoholnih pijač ob eni priložnosti, sledijo pa mu vprašanja o pitju alkoholnih pijač ob ponedeljkih, skupnem dohodku gospodinjstva in pitju alkoholnih pijač on petkih. Socio-demografska vprašanja, s pomočjo katerih bomo v nadaljevanju opisali značilnosti nerespondentov na anketno vprašanje nimajo manjkajočih vrednosti.

Tabela 4.7: Neodgovori spremenljivke v primeru izbranih vprašanj v EHIS 2007.

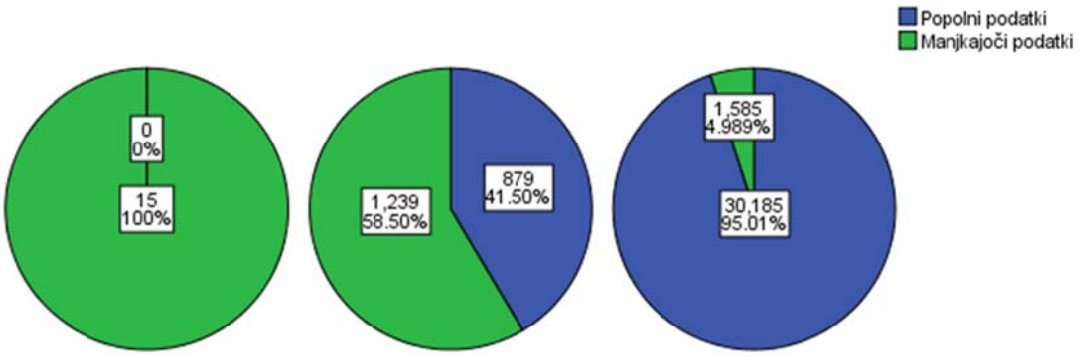
Izbrana anketna vprašanja uporabljena v primeru analize neodgovorov spremenljivke	**n**	**Neodgovor spremenljivke**	
		n	**%**
Kakšno je vaše splošno zdravstveno stanje? (ca1)	2116	2	0,1
Ali imate kakšno dolgotrajno bolezen ali dolgotrajno zdravstveno težavo? (ca2)	2113	5	0,2
V kolikšni meri ste zadnjih 6 mesecih ali dlje ovirani zaradi zdravstvenih težav pri običajnih aktivnostih? (ca3)	2112	6	0,3
Ali imate oz. ste že kdaj imeli katero izmed naslednjih bolezenskih stanj? Kronični bronhitis. (cb1_b)	2115	3	0,1
Ali imate oz. ste že kdaj imeli katero izmed naslednjih bolezenskih stanj? Srčni infarkt. (cb1_c)	2116	2	0,1
Ali imate oz. ste že kdaj imeli katero izmed naslednjih bolezenskih stanj? Bolečine v križu. (cb1_i)	2116	2	0,1
Ali imate oz. ste že kdaj imeli katero izmed naslednjih bolezenskih stanj? Sladkorna bolezen. (cb1_k)	2115	3	0,1
Ali imate oz. ste že kdaj imeli katero izmed naslednjih bolezenskih stanj? Rak (maligni tumor, vključena tudi levkemija in limfom). (cb1_o)	2116	2	0,1
Kako pogosto ste v zadnjih 12 mesecih pili kakršnekoli alkoholne pijače? (ej10)	2113	5	0,2
Koliko pijač, ki vsebujejo alkohol ste popili vsak dan v tipičnem tednu, kadar ste pili alkoholne pijače? Ponedeljek. (al02ap)	1978	140	32,9
Koliko pijač, ki vsebujejo alkohol ste popili vsak dan v tipičnem tednu, kadar ste pili alkoholne pijače? Petek. (al02ep)	2027	91	21,4
Kako pogosto ste v zadnjih 12 mesecih popili 6 ali več meric pijače ob eni priložnosti? (al03p)	1363	755	46,2
Ali ste v zadnjih 12 mesecih uživali kanabis? (cn02)	2112	6	0,3
Število otrok v gospodinjstvu, starih štiri leta ali manj. (hhsize1)	2100	18	0,8
Skupen dohodek gospodinjstva – združena spremenljivka. (doh_gosp)	1572	546	25,8

Vir: EHIS 2007.

Spodnja slika prikazuje splošen pregled neodgovora spremenljivke. Opazimo lahko naslednje:

- vsaj en neodgovor spremenljivke se pojavlja pri vseh izbranih ključnih anketnih vprašanj;
- vsaj en neodgovor spremenljivke ima 58,5 % izmed vseh 2118 respondentov;
- 1585 vrednosti (respondenti x anketna vprašanja) izmed vseh 31770 vrednosti, je manjkajočih.

Slika 4.5: Splošen pregled neodgovora spremenljivke v EHIS.

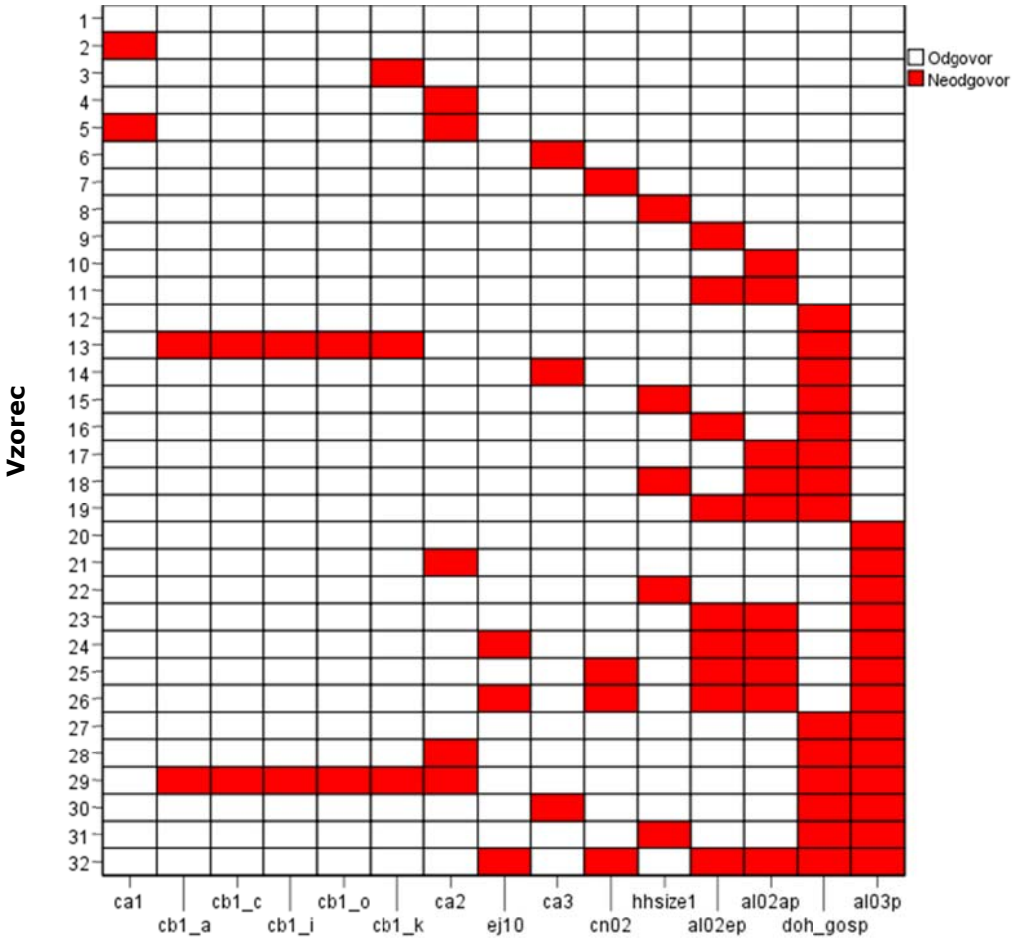


Slika v nadaljevanju besedila prikazuje vzorec neodgovora spremenljivke glede na izbrana anketna vprašanja. Vsak vzorec neodgovora spremenljivke ustreza skupini respondentov s popolnimi podatki (odgovori) in/ ali neodgovori npr.:

- Vzorec 10 predstavlja respondente, ki so niso odgovorili na vprašanje o popitih mericah alkoholnih pijač ob ponedeljkih;
- Vzorec 16 respondente, ki niso odgovorili na vprašanji o popitih mericah alkoholnih pijač ob petkih in skupnem dohodku gospodinjstva;
- Vzorec 13 respondente, ki niso odgovorili na vprašanja o boleznih in bolezenskih stanjih ter vprašanje o skupnem dohodku gospodinjstva;

- Vzorec 6 respondente, ki niso odgovorili na vprašanje o dalj časa trajajoči oviranosti pri vsakdanjih aktivnostih, zaradi zdravstvenih težav;
- Vzorec 32 predstavlja respondente, ki niso želeli odgovoriti na vsa vprašanja o pitju alkoholnih pijač, vprašanje o uživanju marihuane ter vprašanje o skupnem dohodku gospodinjstva;
- Vzorec 1 pa respondente, ki so odgovorili na vsa izbrana anketna vprašanja.

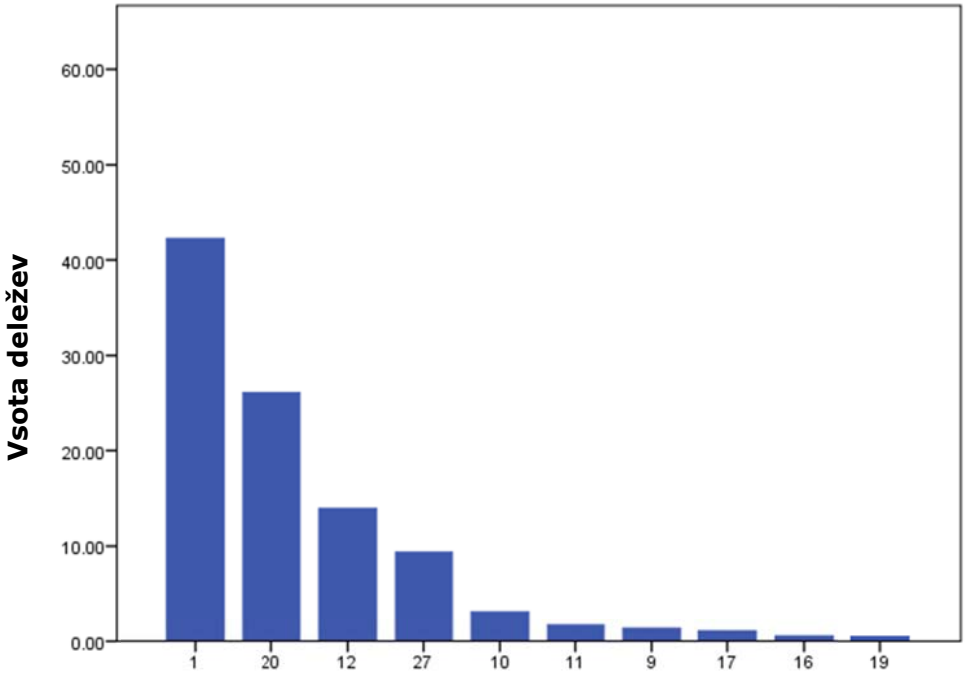
Slika 4.6: Vzorec neodgovora spremenljivke v EHIS.



Poglejmo si še 10 najpogosteje pojavljajočih se vzorcev neodgovorov spremenljivke v izbranem setu podatkov. Opazimo lahko, da dobrih 40 % vseh respondentov pripada vzorcu 1 (respondenti, ki so odgovorili na vsa

anketna vprašanja), dobrih 25 % vzorcu 20 z neodgovori na vprašanje o popitih vsaj šestih mericah alkoholnih pijač ob eni priložnosti, približno 15 % vzorcu 12 z neodgovori na vprašanje o skupnem dohodku gospodinjstva, 10 % pa vzorcu 27 z neodgovori na vprašanji o skupnem dohodku gospodinjstva in popitih vsaj šestih mericah alkoholnih pijač ob eni priložnosti. Ostalim vzorcem pripada manj kot 10 % vseh respondentov.

Slika 4.7: Najpogosteje pojavljajoči vzorci neodgovora spremenljivke v podatkih EHIS.



4.4 Socio-demografske značilnosti nerespondentov

V pričujočem poglavju bodo zgolj za ilustracijo na primeru »Ankete o zdravju in zdravstvenem varstvu« najprej prikazali obravnavane socio-demografske značilnosti nerespondentov na anketo (neodgovor enote), temelječ na socio-demografskih podatkih izbranih oseb iz vzorca. Nato pa še socio-demografske značilnosti nerespondentov na posamezno na anketno vprašanje (neodgovor spremenljivke).

Značilnosti nerespondentov na anketo (neodgovor enote)

Na anketo ni odgovarjalo 1269 oseb, med katerimi z 202 izbranimi osebami anketar ni uspel vzpostaviti stika, 799 oseb je zavrnilo sodelovanje, 268 oseb pa je bilo zaradi različnih razlogov, kot so npr. neobstoj stanovanja na naslovu izbrane osebe, bivanje osebe v tujini, selitev na neznan naslov, selitev v skupinsko gospodinjstvo, smrt osebe ipd., označenih kot neustreznih.

Tabela 4.8: Končni status ankete.

	n	% med nerespondenti	% med vsemi izbranimi osebami
Brez stika (NC[4])	202	15,9	6,0
Nesodelovanje (R[5])	799	62,9	23,6
Neustreznost	268	21,1	7,9
Skupaj	**1269**	**100,0**	**37,5**
Izvedene ankete (I+P)[6]	2118		62,5
Skupaj	**3387**		**100**
Skupaj ustrezni	**3119**		

Vir: EHIS 2007.

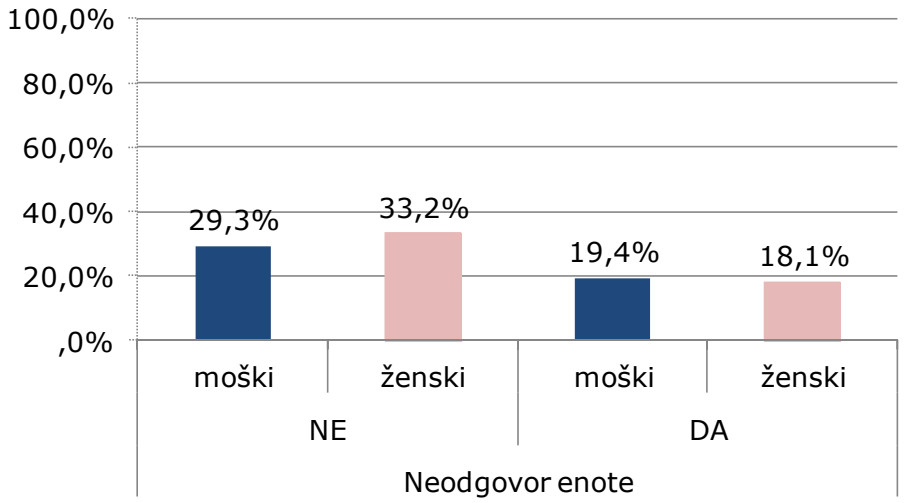
Dosežene stopnje odgovora, sodelovanja, zavračanja in stika v anketi so po metodologiji AAPOR (2011) zgolj za ilustracijo izračunane v Prilogi 1. Stopnja odgovora v anketi je bila 68 %, stopnja sodelovanja 73 %, stopnja zavračanja 26 % in stopnja stika 94 %.

Sliki na naslednji strani prikazujeta deleže neodgovora enote glede na njihov spol in starost. Med vsemi osebami zbranimi vzorec je med neodgovori enote več žensk, medtem ko je med tistimi, ki so odgovarjali

[4] Non-Contact
[5] Refusal
[6] Delne (P) + v celoti izvedene (I) ankete

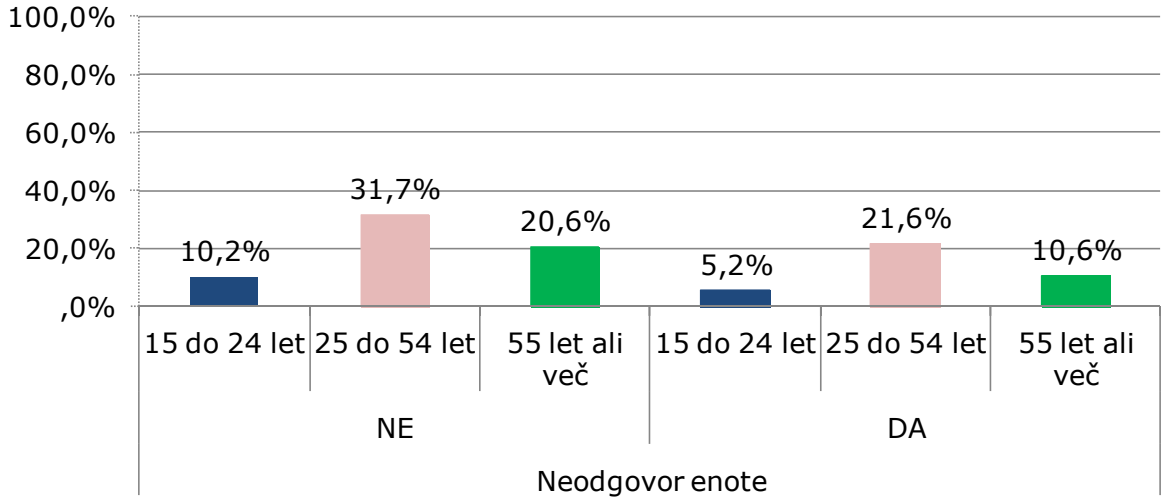
na anketo nekoliko več moških. Med vsemi osebami zbranimi v vzorec pa je tako med neodgovori enote, kakor tudi med tistimi, ki so odgovarjali na anketo največ posameznikov starih od 25 do 54 let.

Slika 4.8: Delež neodgovora enote po spolu med vsemi osebami zbranimi v vzorec.



Vir: EHIS 2007.

Slika 4.9: Delež neodgovora enote po starostnih skupinah med vsemi osebami zbranimi v vzorec.



Vir: EHIS 2007.

Nerespondenti, ki so zavračali sodelovanje v anketi, so bili v primeru, da so to seveda želeli, vprašani tri kratka vprašanja o njihovem zdravstvenem stanju.

Tabela 4.9: Kakšno je vaše splošno zdravstveno stanje? Nerespondenti proti respondentom.

	n	% med vsemi	% med veljavnimi neodgovori enote (neuteženi podatki)	% med veljavnimi neodgovori enote (uteženi podatki[7])	% med odgovori enote (uteženi podatki)
Zelo dobro.	42	3,3	20,3	18,9	21,4
Dobro.	82	6,5	39,6	37,4	40,4
Srednje.	49	3,9	23,7	25,1	27,1
Slabo.	24	1,9	11,6	12,7	8,7
Zelo slabo.	10	0,8	4,8	5,9	2,4
Skupaj	**207**	**16,3**	**100,0**	**100,0**	**100,0**
Ne vem.	1	0,1			
Ne želim odgovoriti.	1060	83,6			
Skupaj manjkajoči	***1061***	***83,7***			
Skupaj	**1268**	**100,0**			

Vir: EHIS 2007.

Na vprašanje o njihovem splošnem zdravstvenem je odgovarjalo 207 siceršnjih nerespondentov, med katerimi jih 56,3 % meni, da je njihovo splošno zdravstveno stanje vsaj dobro, 25,1 % jih meni, da je srednje dobro, 18,6 % pa da je slabo ali zelo slabo. Na drugi strani 61,8 % respondentov meni, da je njihovo splošno zdravstveno stanje vsaj dobro, 27,1 % jih meni, da je srednje dobro, 11,1 % pa jih meni da je njihovo splošno zdravstveno stanje slabo ali zelo slabo.

[7] Utež izračunana s pomočjo post-stratifikacije med vsemi neodgovori enote.

Tabela 4.10: Ali imate kakšno dolgotrajno bolezen ali dolgotrajno zdravstveno težavo? Nerespondenti proti respondentom.

	n	%	% med veljavnimi neodgovori enote (neuteženi podatki)	% med veljavnimi neodgovori enote (uteženi podatki[8])	% med odgovori enote (uteženi podatki)
Da.	59	4,7	29,9	32,8	37,6
Ne.	138	10,9	70,1	67,2	62,4
Skupaj	**197**	**15,5**	**100,0**	**100,0**	**100,0**
Ne želim odgovoriti.	1071	1,1			
Skupaj manjkajoči	***1071***	***84,5***			
Skupaj	**1268**	**100,0**			

Vir: EHIS 2007.

Na vprašanje o tem ali imajo kakšno dolgotrajno bolezen ali dolgotrajno zdravstveno težavo je odgovarjalo 197 siceršnjih nerespondentov, med katerimi jih ima 32,8 % dolgotrajno bolezen ali dolgotrajno zdravstveno težavo. Na drugi strani ima dolgotrajno bolezen ali dolgotrajno zdravstveno težavo 37,6 % respondentov.

Na vprašanje o tem, v kolikšni meri so bili v zadnjih šestih mesecih ali dlje ovirani zaradi zdravstvenih težav pri običajnih aktivnostih, je odgovarjalo 197 siceršnjih nerespondentov, med katerimi jih je bilo 13,4 % zelo oviranih, 24,5 % zmerno oviranih, 62,2 % pa ni bilo oviranih zaradi zdravstvenih težav. Na drugi strani je bilo med respondenti 9,9 % zelo oviranih, 26,4 % zmerno oviranih, 63,7 % pa ni bilo oviranih zaradi zdravstvenih težav.

[8] Utež izračunana s pomočjo post-stratifikacije med vsemi neodgovori enote.

Tabela 4.11: Oviranost zaradi zdravstvenih težav pri običajnih aktivnostih. Nerespondenti proti respondentom.

	n	%	% med veljavnimi neodgovori enote (neuteženi podatki)	% med veljavnimi neodgovori enote (uteženi podatki[9])	% med odgovori enote (uteženi podatki)
Zelo oviran.	24	1,9	12,2	13,4	9,9
Zmerno oviran.	45	3,5	22,8	24,5	26,4
Sploh nisem oviran.	128	10,1	65,0	62,2	63,7
Skupaj	**197**	**15,5**	**100,0**	**100,0**	**100,0**
Ne vem.	1	0,1			
Ne želim odgovoriti.	1071	1,0			
Skupaj manjkajoči	***1071***	***84,5***			
Skupaj	**1268**	**100,0**			

Vir: EHIS 2007.

Za potrebe analize razlik med respondenti in nerespondenti, je bila s pomočjo post-stratifikacije izračunana utež zaradi neodgovora enote. Za izračun uteži so bili uporabljeni populacijski totali oz. populacijske margine predstavljene v prilogi 2 (tabele 7.4 do 7.7).

Na podlagi hi-kvadrat testa (χ^2) ugotavljamo, da med respondenti in nerespondenti, ki so odgovorili na tri kratka vprašanja ni statistično značilnih razlik v splošnem zdravstvenem stanju (p=0,286), dolgotrajnih boleznih oz. dolgotrajnih zdravstvenih težavah (p=0,440) in dalj časa trajajoči oviranosti zaradi zdravstvenih težav (p=0,544). Nerespondenti se torej od respondentov statistično značilno ne razlikujejo.

[9] Utež izračunana s pomočjo post-stratifikacije med vsemi neodgovori enote.

Značilnosti nerespondentov v primeru neodgovora spremenljivke

V nadaljevanju bodo prikazane socio-demografske značilnosti nerespondentov v primeru neodgovora spremenljivke v »Anketi o zdravju in zdravstvenem varstvu«. Socio-demografske značilnosti nerespondentov nam bodo služile kot pojasnjevalne spremenljivke. Čeprav je kognitivni proces neodgovora spremenljivke v teoriji pomembnejši kot socio-demografske značilnosti nerespondentov, pa so spremenljivke, s katerimi opisujemo kognitivno vedenje, redkokdaj dostopne raziskovalcem. Socio-demografske spremenljivke pa so ponavadi zadovoljiv kazalnik osnovnih kognitivnih procesov. Tako je npr. višja izobrazba kazalec višje kognitivne sposobnosti, boljšega razumevanja zastavljenega vprašanja ter širšega znanja in zanimanja. Tudi spol je lahko pomemben kazalnik zanimanja za posamezno vsebinsko tematiko, saj se ženske npr. bolj zanimajo za družino in otroke, medtem ko se moški bolj zanimajo za politiko (Kveder 2005).

Neodvisne spremenljivke, na podlagi katerih bomo določili socio-demografske značilnosti nerespondentov v primeru neodgovora spremenljivke v »Anketi o zdravju in zdravstvenem varstvu«, so naslednje:

a) spol,
b) starost,
c) izobrazba,
d) zakonski stan,
e) poklic,
f) gospodinjstvo z otroki,
g) velikost naselja in
h) način anketiranja (samoanketiranje v primerjavi z anketiranjem s pomočjo anketarja z vprašalnikom na papirju).

Starost anketiranca smo prekodirali v tri kategorije, in sicer:

1. osebe stare od 15 do 24 let,
2. osebe stare od 25 do 54 let in
3. osebe starejše od 55 let.

Doseženo stopnjo izobrazbe anketiranca smo skladno s Klasifikacijskim sistemom izobraževanja in usposabljanja prekodirali v tri kategorije, in sicer dosežena:

1. osnovnošolska izobrazba ali manj (združene kategorije: brez šolske izobrazbe, nepopolna osnovna izobrazba osnovna izobrazba),
2. srednješolska izobrazba (združene kategorije: nižja ali srednja poklicna izobrazba, srednja strokovna izobrazba, srednja splošna izobrazba),
3. višješolska, visokošolska izobrazba (združene kategorije: višja strokovna izobrazba, višješolska izobrazba, visoka strokovna izobrazba, visoka univerzitetna izobrazba, specialistična povisokošolska izobrazba, magisterij, doktorat).

Spremenljivko *formalni zakonski stan* smo prekodirali v tri kategorije:

1. samski, nikoli poročen
2. poročen
3. drugo (združeni kategoriji: ovdovel, ločen).

Spremenljivko *trenutni zaposlitveni status* smo prekodirali v pet kategorij:

1. zaposlen
2. brezposeln
3. šolajoč
4. upokojen
5. drugo (združene kategorije: trajno invalidna oseba, obvezno služenje v vojski ali skupnosti, gospodinja, drugo).

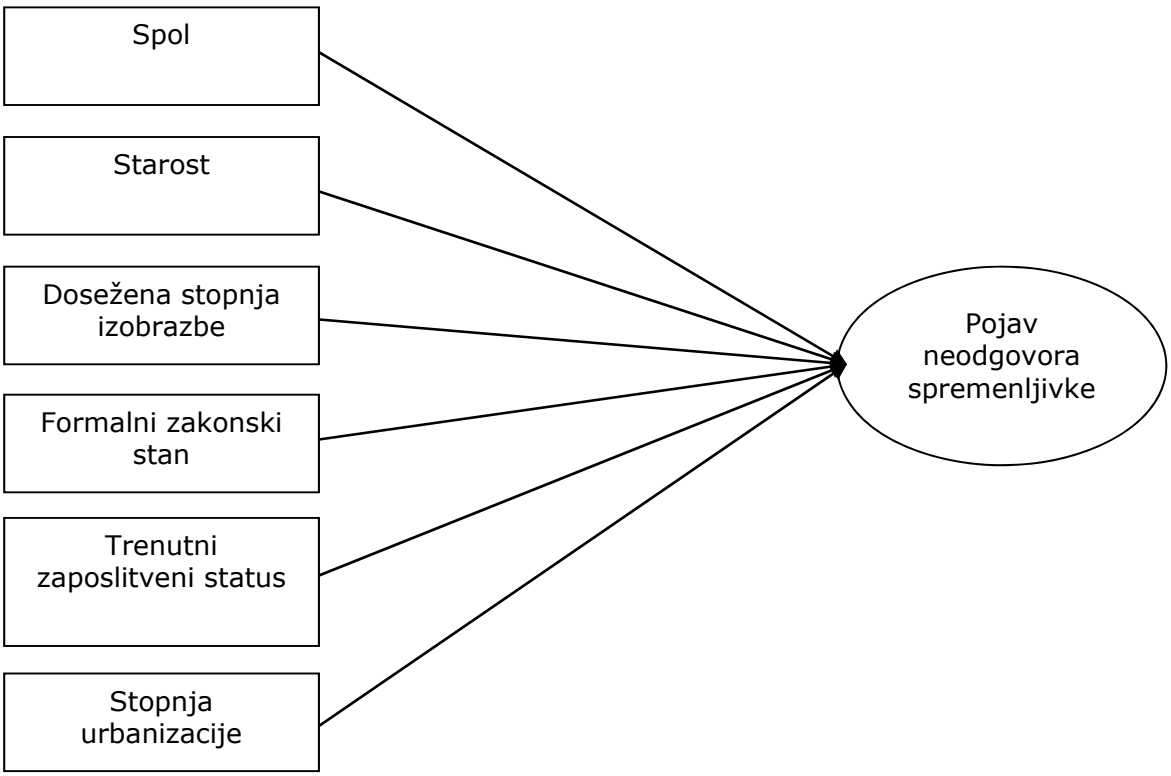
Spremenljivko *stopnja urbanizacije* smo dobili iz vzorca, njene vrednosti pa pomenijo naslednje:

1. gosto poseljena območja,
2. srednje poseljena območja,
3. redko poseljena območja.

Spremenljivko *pojav neodgovora spremenljivke* smo izračunali tako, da smo vsaki izbrani spremenljivki pripisali vrednost 0, če je oseba odgovorila na izbrano vprašanje in 1, če ni (neodgovor spremenljivke). Nato smo neodgovore spremenljivke sešteli v skupno število neodgovorov spremenljivke, novo izračunano spremenljivko pa nazadnje prekodirali vrednost 0, če je oseba odgovorila na vsa izbrana anketna vprašanja in 1, če oseba ni odgovorila vsaj na enega izmed izbranih vprašanj. Vrednost 1, torej predstavlja vsaj en neodgovor spremenljivke, vrednost 0 pa odgovor na vsa izbrana anketna vprašanja.

Vpliv izbranih socio-demografskih značilnosti na pojav neodgovora spremenljivke bomo v nadaljevanju preverjali s pomočjo logistične regresije. Spodnja slika predstavlja osnovne socio-demografske dejavnike, za katere na podlagi ugotovitev preteklih raziskav menimo, da bi lahko imeli vpliv na pojav neodgovorov.

Slika 4.10: Raziskovalni model socio-demografskih dejavnikov, za katere se predvideva, da vplivajo na pojav neodgovora spremenljivke.



Ker je en izmed zahtevanih pogojev za izvedbo logistične regresije neobstoj kolinearnosti med opazovanimi pojasnjevalnimi spremenljivkami, smo najprej izvedli test multikolinearnosti (rezultati testa so v prilogi). Glede na faktorje VIF, lastne vrednosti in z majhnimi lastnimi vrednostmi povezane nizke deleže variance ugotavljamo, da odvisnost med opazovanimi pojasnjevalnimi spremenljivkami ne obstaja, zato lahko nadaljujemo z analizo s pomočjo logistične regresije. V model logistične regresije smo vključili vse opazovane pojasnjevalne spremenljivke za katere na podlagi preteklih raziskav predvidevamo, da bi lahko imeli vpliv na pojav neodgovorov. Rezultati logistične regresije so prikazani v spodnji tabeli.

Tabela 4.12: Model logistične regresije (model 1).

	Ocena parametra (β)	Standardna napaka	Wald	SP	Stat. značilnost	Razmerje obetov (exp(β))	95% IZ za razmerje obetov	
							Spodnja meja	Zgornja meja
Spol								
ženski	0,278	0,107	6,797	1	0,009	1,321	1,072	1,629
Starostna skupina								
25 do 54 let	0,185	0,261	0,502	1	0,479	1,203	0,722	2,004
55 let in več	0,186	0,311	0,355	1	0,551	1,204	0,654	2,216
Stopnja urbanizacije								
Srednje poseljena območja	0,216	0,150	2,070	1	0,150	1,241	0,925	1,665
Gosto poseljena območja	0,175	0,148	1,398	1	0,237	1,192	0,891	1,593
Zakonski stan								
Poročen	0,121	0,138	0,760	1	0,383	1,128	0,860	1,479
Drugo	0,132	0,203	0,420	1	0,517	1,141	0,766	1,699
Zaposlitveni status								
Zaposlen	0,402	0,268	2,242	1	0,134	1,495	0,883	2,529
Brezposeln	-0,113	0,310	0,133	1	0,715	0,893	0,486	1,640
Šolajoč	0,609	0,366	2,768	1	0,096	1,839	0,897	3,767
Upokojen	0,053	0,270	0,038	1	0,845	1,054	0,621	1,791
Dosežena stopnja izobrazbe								
Srednješolska	0,209	0,131	2,552	1	0,110	1,232	0,954	1,591
Visokošolska, višješolska	0,430	0,187	5,313	1	0,021	1,538	1,067	2,218
Skupen dohodek gospodinjstva								
Do 650 EUR	-0,017	0,197	0,007	1	0,932	0,983	0,668	1,448
Od 650 do 850 EUR	0,227	0,161	1,984	1	0,159	1,255	0,915	1,721
Od 851 do 1050 EUR	-0,040	0,163	0,059	1	0,808	0,961	0,698	1,323
Od 1051 do 1850 EUR	0,013	0,166	0,006	1	0,937	1,013	0,732	1,403
Konstanta	-1,007	0,399	6,377	1	0,012	0,365		

Model je kot celota statistično značilen (p=0,009), Hosmer-Lemeshow test pa kaže, da se od podatkov ne razlikuje statistično značilno (p=0,132). Z modelom smo uspeli pojasniti le 2,9 %[10] variabilnosti odvisne spremenljivke.

Na podlagi statistične značilnosti Waldove statistike ugotavljamo, da med izbranimi pojasnjevalnimi spremenljivkami na pojav neodgovora spremenljivke, vplivata le spol in dosežena stopnja izobrazbe

[10] Nagelkerke R^2

posameznika. Ženske imajo tako 1,3-krat višje obete za pojav neodgovora spremenljivke, kot moški (95% IZ[11] je od 1,07 do 1,63). Podobno pa imajo posamezniki z doseženo visokošolsko ali višješolsko izobrazbo 1,5-krat višje obete za pojav neodgovora spremenljivke, kot posamezniki z doseženo osnovnošolsko izobrazbo ali manj (95% IZ je od 1,07 do 2,22).

Tudi Kveder (2005) ugotavlja podobno, in sicer v analizi treh različnih anketnih raziskav ugotavlja, da se večji deleži neodgovora spremenljivke pojavljajo pri ženskah, vendar se vpliv spola povezuje tudi s tematiko anketne raziskave in zanimanjem za odgovarjanje na posamezna anketna vprašanja. Njegova analiza vpliva izobrazbe na neodgovor spremenljivke je pokazala ravno nasproten vpliv kot ga ugotavljamo v »Anketi o zdravju in zdravstvenem varstvu«. Ugotavlja namreč, da se stopnje neodgovora spremenljivke višajo z nižanjem izobrazbe posameznika. Njegova analiza je pokazala tudi vpliv starosti na neodgovor spremenljivke, in sicer se višji deleži neodgovora spremenljivke pojavljajo pri starejših anketirancih.

4.5 Reševanje problema neodgovorov

Neodgovor spremenljivke predstavlja problem, ki se pogosto pojavlja v anketnem raziskovanju, še posebej v anketah, povezanih z zdravjem, in družboslovnih anketah. Spremenljivke, ki raziskovalce pogosto še posebej zanimajo, kot so npr. dohodek, mnenja, odnos in prepričanja posameznikov ter druga občutljiva vprašanja, so ponavadi še posebej nagnjena k neodgovorom (Durrant 2005). Četudi neodgovora spremenljivke ne moremo nikoli v celoti preprečiti, pa ga lahko znatno omejimo. Preprečevanje pojava neodgovora spremenljivke in vstavljanje manjkajočih vrednosti sta med seboj močno povezana v procesu

[11] Interval zaupanja

reševanja problema neodgovora spremenljivke, vendar pa je preprečevanje pojava vedno prvi korak, ko gre za reševanje problema neodgovora spremenljivke, saj posledično vodi v manj vstavljanja manjkajočih podatkov v obravnavani set podatkov in nenazadnje imajo posledično raziskovalci na voljo večji nabor podatkov kot osnovo za izvedbo vstavljanja manjkajočih podatkov (De Leeuw in drugi 2003).

Stopnje neodgovora spremenljivke lahko rešujemo na različne načine. V anketnih raziskovanjih med otroki in mladimi s samoanketiranjem z vprašalnikom na papirju lahko problem neodgovora spremenljivke rešujemo že v fazi priprave anketnega vprašalnika, in sicer se lahko visokim stopnjam neodgovora spremenljivke izognemo že s spremembo vrstnega reda vprašanj, npr. vprašanja iz zadnjega dela vprašalnika, ki so za nas morda bolj pomembna, prerazporedimo preko celotnega vprašalnika. S tem prerazporedimo stopnje neodgovorov po posameznih vprašanjih, kar posledično pomeni, da niso skoncentrirane pri vprašanjih v zadnjem delu vprašalnika. V anketnih raziskovanjih med otroki in mladimi k višjim stopnjam odgovora spremenljivke do neke mere pripomore tudi obsežno uvodno besedilo k posameznemu vprašanju. Tudi znižanje števila možnih odgovorov lahko pripomore k nižanju neodgovora spremenljivke (Borgers in Hox 2001).

Vpliv na stopnje neodgovora spremenljivke ima tudi uporaba vejitve anketnih vprašanj, v primeru katere namesto postavitve enega anketnega vprašanja z večjim številom možnih kategorij odgovorov, anketirancu zastavimo dve povezani vprašanji z manjšim številom možnih kategorij odgovorov. Raziskave sicer ugotavljajo, da dodatna navodila glede vejitev pomembno zvišujejo deleže neodgovora spremenljivke. Na neodgovore v primeru vejitev pomembno vpliva tudi starost anketiranca, in sicer se višji deleži pojavljajo v primeru starejših anketirancev. Raziskovalci morajo tako že v fazi priprave anketnega vprašalnika ugotoviti, katere oblike

vprašanj so primernejše za specifične populacijske skupine (Messmer in Seymour 1982).

Poleg načinov za reševanje problemov zaradi pojava neodgovora spremenljivke, ki jih uporabljamo že v fazi priprave anketnega vprašalnika, se raziskovalci pogosto poslužujejo tudi metod, s katerimi dopolnjujejo oziroma vstavljajo manjkajoče podatke.
Izbor primerne metode ravnanja z manjkajočimi podatki v toku analiziranja podatkov je eden izmed izzivov, s katerimi se soočajo raziskovalci. Pogosto so manjkajoče vrednosti v podatkovnih zbirkah enostavno prezrte, namesto da bi bile obravnavane z zanesljivo metodo vstavljanja manjkajočih podatkov (Saunders in drugi 2006).

Rässler in Riphahn (2006) ugotavljata, da neuporaba metod uteževanja in vstavljanja manjkajočih podatkov lahko vodi k precejšni pristranskosti dobljenih ocen. Izvajanje enkratnega vstavljanja manjkajočih podatkov pa je odvisno od tega, ali obstaja vzorec manjkajočih podatkov in od informacij o tem, ali podatki manjkajo (popolnoma) slučajno. S pomočjo večkratnih vstavljanj ponavadi dobimo najboljše pokritje prave vrednosti, odsevajo pa tudi obstoječe vzorce povezanosti.

Vstavljanje manjkajočih podatkov je sicer široko uporabljana metoda v anketnih raziskovanjih, vendar pa lahko privede tudi do napačnih sklepanj, npr. neveljavnih intervalov zaupanja (Deville in Särndal 1994). Pomembno je, da za potrebe statističnega vstavljanja neodgovora spremenljivke ločimo neodgovore, ki izvirajo iz nenamernih preskokov posameznih anketnih vprašanj, in neodgovorov, ki izvirajo iz posameznikove nepripravljenosti razkritja občutljivih informacij (De Leeuw in drugi 2003).

Obstajajo različne metode, s katerimi vstavljamo manjkajoče podatke (Vehovar 2007):

(a) *Vstavljanje aritmetične sredine* je najbolj enostavna oblika vstavljanja manjkajočih podatkov. Enostavnost takega pristopa pa je pogojno – zgolj za ocenjevanje aritmetične sredine, ne pa tudi za ocenjevanje variance – upravičena le, če velja za manjkajoče podatke lastnost MCAR. Le v tem primeru imajo namreč manjkajoče vrednosti enako aritmetično sredino kot vrednosti respondentov. Vstavljanje aritmetične sredine znotraj razredov uteževanja pa daje v smislu ocenjene aritmetične sredine skoraj enake rezultate kot uteževanje.

(b) *Vstavljanje z regresijsko metodo* je način, pri katerem namesto manjkajoče vrednosti vstavimo oceno na osnovi regresijske spremenljivke *Y* na ostale spremenljivke, ki so na odgovarjajočih podatkih opazovane popolno. Vstavljanje z regresijsko metodo lahko omogoča bistveno boljše rezultate kot vstavljanje povprečne vrednosti. Kot točkovna ocena manjkajoče vrednosti je zgornja ocena celo optimalna, saj pod predpostavko MCAR minimizira srednjo kvadratno napako (MSE – angl. »mean squared error«). Vstavljanje na osnovi regresijske funkcije je primerno tudi pri predpostavki MAR, pri čemer mora biti povezanost spremenljivk linearna.

(c) *Vstavljanje vrednosti respondentov.* Namesto manjkajoče vrednosti pogosto vstavljamo razpoložljive vrednosti respondentov. Take metode se pogosto imenujejo tudi metode »hot-deck«. Privlačnost takega pristopa predstavlja predvsem dejstvo, da ohranjamo porazdelitev spremenljivke med respondenti, kar pri vstavljanju povprečne vrednosti ni bilo

mogoče. Pri vseh enostavnih oblikah tovrstnega vstavljanja vedno postavljamo za manjkajoče vrednosti pogoj MCAR. Postopki se med seboj razlikujejo predvsem v pogledu izbire nadomestnih vrednosti, za kar je več možnosti:

- *Enostavna slučajna izbira respondentov*. V primeru enostavnega slučajnega izbiranja nadomestne enote s ponavljanjem iz množice respondentov se izkaže, da je ocena aritmetične sredine nepristranska.

- *Zaporedno vstavljanje* (angl. »segmential hot-deck«). Namesto manjkajoče vrednosti lahko vstavimo kar vrednosti predhodnega respondenta iz zbranega seta podatkov, kar močno poenostavlja »tehnologijo« vstavljanja. Povečanje variance je občutnejše kot v primeru slučajnega izbiranja respondentov, saj je intuitivno jasno, da je proces izbire v tem primeru nekoliko »manj slučajen«. V primeru določene zakonitosti, npr. povezanosti sosednjih elementov, v vrstnem redu enot, pa se varianca seveda poveča še za vpliv korelacije med vrednostmi sosednjih členov.

- *Kompleksne metode vstavljanja*:
  - *Vstavljanje vrednosti respondentov znotraj razredov*. Nadomestne vrednosti izbiramo znotraj razredov na osnovi zgoraj opisanih postopkov. Tako dobljene ocene združimo v skupno oceno podobno kot v primeru uteževanja. Dodati velja, da se v pogojih stratificiranega vzorčenja oziroma vzorčenja v skupinah povečanje variance, ki nastane zaradi zgoraj opisanih postopkov vstavljanja vrednosti respondentov, obnaša enako kot siceršnja varianca v stratumih oziroma skupinah. Doseči moramo le, da je

postopek med stratumi oziroma skupinami popolnoma neodvisen.

- *Vstavljanje najbližje enote*. Alternativen pristop je vpeljava določene metrike, ki temelji na znanih vrednostih kovariat. Na tej osnovi izračunavamo razdalje med enotami in določeni enoti namesto manjkajoče vrednosti vstavimo vrednosti njej najbližje enote. Ker pa so tako vstavljene vrednosti razmeroma zapletene funkcije, je v takem primeru razmeroma težko izpeljati ocene vzorčne variance.
- *Vstavljanje na osnovi modela*. Namesto manjkajočih vrednosti lahko vstavimo tudi vrednosti, ki jih izračunamo na osnovi določenega modela (npr. na osnovi regresijskega modela).
- *Vstavljanje vrednosti na osnovi zunanje informacije*. Vrednosti lahko vstavljamo tudi na osnovi zunanje informacije. Gre npr. za podatke iz popisa, druge ankete ali pa ekspertne ocene. Seveda to niso vzorčni podatki in morebitnih napak, ki jih takšni postopki vsebujejo, ne moremo kontrolirati, zato tudi težko ocenimo varianco.
- *Vstavljanje nadomestne vrednosti* (rezervne enote so poseben primer vstavljanja).
- *Večkratno vstavljanje*. Vrednosti lahko vstavimo tudi večkrat, kar v mnogočem izboljšuje ocene.

Največja prednost vstavljenih podatkov je v tem, da po opravljenem postopku vstavljanja dobimo popolne podatke, zaradi česar lahko v nadaljevanju izvajamo statistične analize na popolnem setu podatkov. Naslednja velika prednost vstavljanja manjkajočih podatkov je tudi v tem, da se lahko postopek vstavljanja podatkov pogosto izvede samo enkrat, v primeru, ko raziskovalec dobro pozna vzroke za neodgovore. To je mogoče samo takrat, ko ima raziskovalec na voljo več informacij, ki mu služijo v pomoč v postopku vstavljanja podatkov, npr. iz administrativnih

virov (Rubin 2004, 11–12). Pa vendar, kljub prednostim, ki jih vstavljanje prinaša, ostaja tudi velik problem. Namreč to, da manjkajoče vrednosti niso znane in z vstavljanjem dopolnjen set podatkov obravnavamo kot popoln, lahko povzroči pristranske ocene podatkov, saj variabilnosti zaradi neznanih manjkajočih vrednosti ne poznamo (Rubin 2004, 12–13).

Podobno kot enkratno vstavljanje manjkajočih podatkov tudi večkratno vstavljanje predstavlja način, s pomočjo katerega dopolnjujemo manjkajoče podatke. V primeru večkratnih vstavljanj za vsako manjkajočo vrednost vstavimo več vrednosti, npr. *m* vrednosti, namesto ene. Z večkratnimi vstavljanji tako dobimo *m* različnih popolnih setov podatkov, na katerih nato izvajamo različne statistične analize. Podobno kot v primeru enkratnega vstavljanja tudi z uporabo večkratnih vstavljanj po opravljenem postopku vstavljanja dobimo popoln set podatkov, zaradi česar lahko v nadaljevanju izvajamo statistične analize na popolnih podatkih. Raziskovalec, ki dobro pozna vzroke za neodgovore spremenljivke v danih podatkih pa lahko v postopku vstavljanja uporabi svoje znanje, pri čemer lahko ob uporabi večkratnih vstavljanj uporabi tudi različne vrednosti v primerih, ko pričakovana vstavljena vrednost ni zanesljiva ter na tak način v večkratnih vstavljanjih uporabi različne pričakovane vrednosti neke manjkajoče vrednosti (Rubin 2004, 15–16).

Poleg zgoraj navedenih dveh prednosti tako enkratnih kakor tudi večkratnih vstavljanj manjkajočih podatkov imajo večkratna vstavljanja še tri pomembne prednosti, in sicer (Rubin 2004, 16):

a) kadar so vstavljanja slučajno porazdeljena v poskusu porazdelitve podatkov, večkratno vstavljanje povečuje učinkovitost ocen,
b) kadar večkratno vstavljanje predstavlja ponavljajoč slučajni vzorec v modelu neodgovora, pridemo do veljavnih sklepanj enostavno z združitvijo popolnih setov podatkov,
c) generiranje slučajnih vstavljanj znotraj različnih modelov neodgovorov omogoča enostavno proučevanje občutljivosti sklepanj

v različnih modelih neodgovorov s preprosto uporabo ponavljajočih statističnih metod na popolnih setih podatkov.

Vendar pa imajo večkratna vstavljanja v primerjavi z enkratnim vstavljanjem manjkajočih podatkov tudi nekaj pomanjkljivosti, in sicer (Rubin 2004, 17–18):

a) izvedba večkratnih vstavljanj pomeni bistveno več dela kot izvedba enkratnega vstavljanja,
b) za hrambo setov podatkov, pridobljenih z večkratnimi vstavljanji potrebujemo več prostora,
c) za analizo večkrat vstavljenih popolnih setov podatkov je potrebnega tudi več analitičnega dela.

V praksi se v anketnih raziskovanjih za vstavljanje podatkov zaradi neodgovora spremenljivke najpogosteje uporablja metoda enkratnega vstavljanja respondentov (metoda »hot-deck«), pri kateri se vsako manjkajočo vrednost nadomesti z opazovano vrednostjo podobne enote oziroma vstavi vrednost podobnega respondenta. Metoda vstavljanja vrednosti podobnih respondentov ne temelji na prileganju modela spremenljivke, ki jo vstavljamo, temelječ na parametričnem modelu tako kot vstavljanje, saj lahko vstavimo le verjetne vrednosti opazovane spremenljivke, ki izvirajo iz opazovanih vrednosti »darovalcev« (Andridge in Little 2009).

Problem pojava neodgovora spremenljivke lahko torej rešujemo z vstavljanjem manjkajočih podatkov, lahko pa se manjkajočim podatkom enostavno pripišejo kode, ki pomenijo manjkajoče podatke in se kot take izločijo iz nadaljnjih analiz zbranih podatkov (Winglee in drugi 2001).

Widaman (2006) predlaga, da se, kadar je pojav neodgovora spremenljivke zgolj slučajen in so stopnje neodgovora spremenljivke relativno nizke (npr. 1 % do 2 %), uporablja enkratno vstavljanje

manjkajočih podatkov. Kadar pa so stopnje neodgovora spremenljivke višje in spremenljivke zaradi visokih deležev neodgovora ne morejo biti vključene v analize, je priporočljivo večkratno vstavljanje manjkajočih podatkov.

Zgolj za ilustracijo si v nadaljevanju poglejmo način reševanja problema neodgovora spremenljivke saj uteževanje zbranih podatkov zaradi neodgovora enote, predstavlja enega izmed načinov reševanja problema neodgovorov v splošnem.

Uteževanje podatkov zaradi neodgovora enote

Neodgovori v anketnih raziskovanjih lahko pomembno vplivajo na kakovost ocen, kadar se nerespondenti bistveno razlikujejo od respondentov glede na tematiko raziskave (Schouten 2007).

Vpliv neodgovorov v anketnih raziskovanjih lahko zmanjšamo poleg zgoraj opisanih metod tudi s pomočjo metod uteževanja podatkov. V procesu uteževanja podatkov vsakemu posameznemu anketirancu pripišemo utež, s pomočjo katere dosežemo, da dobljena vzorčna struktura čim bolj odseva dejansko populacijsko strukturo.

Metode uteževanja podatkov temeljijo na t. i. pomožnih informacijah, pridobljenih iz populacijskih baz podatkov, popisov in registrov. Na podlagi tako zbranih pomožnih informacij poznamo npr. starostno in spolno strukturo ciljne populacije in s tem tudi starostno in spolno strukturo vzorca (Schouten 2007).

Na prvi stopnji uteževanja vsaki izbrani vzorčni enoti, ne glede na to, ali gre za respondenta ali nerespondenta, pripišemo osnovno oziroma vzorčno utež, ki predstavlja verjetnost izbora posameznega elementa v vzorec. Na drugi stopnji procesa uteževanja gre ponavadi za uteževanje

zaradi neodgovorov na celotno anketo, pri čemer so t. i. vzorčne uteži izravnane glede na posamezne enote, ki niso sodelovale v raziskavi. Tukaj je treba prepoznati značilnosti respondentov, ki so podobni nerespondentom, pri čemer si pomagamo s pomožnimi spremenljivkami, ki jih poznamo za respondente in za nerespondente (npr. spol, starost, izobrazba ...), nato pa povečamo vzorčno utež na način, da respondenti predstavljajo tudi podobne nerespondente. Glavni namen metod uteževanja je zmanjšanje pristranskosti vzorčnih ocen zaradi neodgovora (Kalton in Flores-Cervantes 2003).

Ena izmed pomembnejših posledic uteževanja podatkov je povečanje vzorčne variance. Ocena povečanja vzorčne variance temelji na znani formuli za izračun koeficienta variacije uteži:

$$CV(w) = \frac{\sqrt{s_w^2}}{\overline{w}} \text{ (Kish 1965)},$$

kjer $CV^2(w)$ pomeni kvadrat koeficienta variacije in izraža razmerje med elementarno varianco uteži in kvadratom aritmetične sredine. Kvadrat koeficienta variacije je vključen v faktor povečanja vzorčne variance VIF (angl. »variance inflation factor«), ki izraža povečanje vzorčne variance uteženega vzorca v primerjavi z vzorčno varianco vzorca enake velikosti, kjer uteži niso bile potrebne:

$$VIF = 1 + CV^2(w).$$

Posledica uteževanja je povečanje vzorčne variance za faktor povečanja vzorčne variance:

$$Var(\overline{y}_W) = Var(\overline{y}) \times VIF \text{ (Vehovar in Zupanič 2007)}.$$

Poznamo različne metode uteževanja (Kalton in Flores-Cervantes 2003):

(a) *Uteževanje celic* oziroma *poststratifikacija*, pri čemer vzorčne uteži prilagodimo tako, da ustrezajo populacijskim totalom, temelječ na posameznih celicah glede na dane pomožne spremenljivke. Slabost post-stratifikacije je v tem, da lahko vodi

k veliki variabilnosti v porazdelitvi uteži, na podlagi česar se lahko močno poveča varianca vzorčnih ocen.

(b) Metoda *raking* uporablja v procesu uteževanja samo robne porazdelitve pomožnih spremenljivk in predstavlja iterativen proces prileganja, pri čemer najprej prisilimo totale vrstic, da se prilegajo populacijskim totalom vrstic, nato pa še totale stolpcev, da se prilegajo populacijskim totalom stolpcev. Postopek ponavljamo tako dolgo, da skonvergira. Ker v raziskovanjih pogosto poznamo le robne porazdelitve pomožnih spremenljivk, je metoda raking pogosteje uporabljena kot poststratifikacija.

(c) *Linearno uteževanje* podobno kot raking prilagaja uteži tako, da robne porazdelitve pomožnih spremenljivk vzorca ustrezajo populacijskim robnim porazdelitvam. Linearno uteževanje je poseben primer splošne regresijske ocene, njegova slabost pa je v tem, da lahko v nekaterih primerih vodi do negativnih uteži.

(d) *Uteževanje GREG* (angl. »generalized regression weighting«) v anketnem raziskovanju izhaja iz standardne regresijske ocene, pri čemer sprememba uteži glede na prilagajanje pomožnim spremenljivkam predstavlja regresijsko oceno. Uteževanje GREG v regresijskem modelu zlahka obsega različne pomožne spremenljivke in vključuje neenake uteži. Čeprav GREG uteži, izhajajo iz regresijskega modela, kjer je odvisna spremenljivka y, te niso odvisne od y, ampak so splošne uteži, ki se lahko uporabljajo za vse analize anketnih podatkov.

(e) *Uteževanje s pomočjo logistične regresije* se uporablja z namenom razvoja prilagoditve uteži za neodgovor. Z modelom logistične regresije predvidevamo verjetnost odgovora v primeru, da vzorec temelji na pomožnih spremenljivkah, prilagoditev uteži za vsakega respondenta pa je enaka obratni vrednosti respondentove napovedane verjetnosti odgovora. Kadar so pomožne spremenljivke kategorične in model ne vsebuje

interakcij, je uteževanje s pomočjo logistične regresije podobno rakingu.

(f) *Mešanica post-stratifikacije in drugih metod* se lahko uporablja, kadar uporaba samo post-stratifikacije pomeni nestabilne uteži, ki se pogosto pojavijo v primeru zelo majhnega števila enot v posamezni celici vzorca. V taki situaciji je kompromisna rešitev uporaba post-stratifikacije pri celicah z večjim številom enot in uporaba rakinga pri celicah z zelo majhnim številom enot.

Pojav pristranskosti ocen zaradi neodgovora se pojavi kot funkcija korelacije nagnjenosti posameznika k odgovoru in lastnosti, ki jih meri raziskovalec. Tudi znotraj iste raziskave imajo lahko različne vzorčne ocene različne pristranskosti zaradi neodgovora (Groves 2006).

Pristranskost ocen merimo z naslednjimi izrazi (Vehovar in Zupanič 2007):

a) $bias(\bar{y}) = \bar{y} - \bar{y}_w$;

b) arbias (absolutna relativna pristranskost) $= \left| \frac{bias(\bar{y}_W)}{\bar{y}_W} \right|$, meri velikost pristranskosti glede na samo oceno;

c) asbias (absolutna standardizirana pristranskost) $= \left| \frac{bias(\bar{y}_W)}{se(\bar{y}_{SRS})} \right|$,

primerja pristranskost zaradi neodgovora s standardno napako ocene, npr. z vzorčno napako.

Družbena zaželenost tematike posameznih vprašanj pomembno vpliva na stopnje neodgovora in z njimi povezano pristranskost ocen. Neodgovor spremenljivke ima tako manjši vpliv na ocene, vezane na družbeno zaželene in nevtralne značilnosti (Sakshaug in drugi 2010).

Uteževanje v primeru Ankete o zdravju in zdravstvenem varstvu

V nadaljevanju bo izračunana utež zaradi neodgovora v primeru Ankete o zdravju in zdravstvenem varstvu. Pred uteževanjem in po njem bodo analizirana tudi tri kratka vprašanja o zdravstvenem stanju, ki so bila zastavljena tudi nerespondentom, na podlagi česar bo izračunano povečanje variance zaradi uteževanje in preverjena koristnost uteževanja.

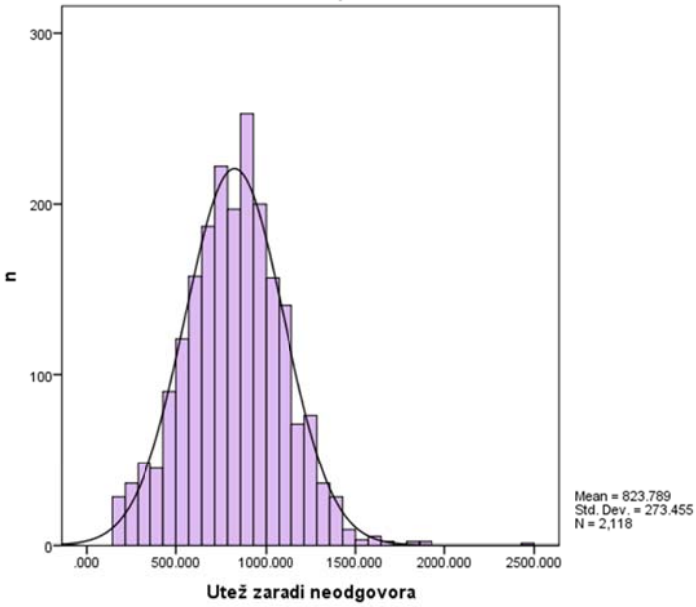
Utež zaradi neodgovora na celotno anketo je bila za vsakega posameznega respondenta izračunana z uporabo programa Blaise, v okviru katerega smo uporabili programski paket Bascula. V Basculi za vse respondente v vzorcu s pomočjo pomožnih informacij izračunamo uteži zaradi neodgovora. Utež zaradi neodgovora enote smo izračunali s pomočjo več pomožnih spremenljivk, pri čemer smo dobljeno vzorčno strukturo prilagajali dejanski populacijski strukturi na podlagi informacij o populacijski strukturi opazovane populacije v Centralnem registru prebivalcev po (uporabljeni populacijski totali so v tabeli 5.4):

- spolu,
- petletnih starostnih skupinah,
- tipu naselja,
- statistični regiji ter
- doseženi stopnji izobrazbe (temelječ na podatkih »Ankete o delovni sili« izvedene leta 2007).

V procesu uteževanja smo uporabili tudi utež verjetnosti izbora v vzorec. Ker uteževanja, zaradi nezadostnih podatkov v večrazsežni kontingenčni tabeli, glede na vse pomožne spremenljivke nismo mogli izvesti s pomočjo metode post-stratifikacije, smo uporabili *linearno uteževanje*, v postopku katerega smo več različnih mejnih vzorčnih totalov, prilagajali mejnim populacijskim totalom. Metoda linearnega uteževanja temelji na splošni regresijski oceni. Za vzorce zadostne velikosti pa je regresijska ocena nepristranska.

Spodnja slika prikazuje končno porazdelitev izračunanih uteži zaradi neodgovorov.

Slika 4.11: Porazdelitev uteži zaradi neodgovora enote v EHIS.



Vir: EHIS 2007.

Na podlagi dobljenih uteži smo najprej izračunali faktor povečanja variance (*vif*), ki predstavlja povečanje vzorčne variance uteženega vzorca v primerjavi z vzorčno varianco vzorca enake velikosti, kjer uteži niso bile potrebne:

$$vif = 1 + cv^2(w) = 1{,}11$$

Vzorčna varianca se zaradi uteževanja poveča za 11 %, torej uteževanje poveča pristranskost.

Koristnost uteževanja bomo preverili v primeru spremenljivk splošno zdravstveno stanje, pitje katerih koli alkoholnih pijač, pitje šest ali več meric ob eni priložnosti in združen skupen dohodek gospodinjstva.

S pomočjo naslednjih formul bomo izračunali srednje kvadratno napako (mse), ki jo izračunamo kot vsoto kvadrata pristranskosti in variance:

$$mse(\bar{y}) = bias^2(\bar{y}) + var(\bar{y})$$

$$mse(\overline{y}_w) = var(\overline{y}) \times vif$$

Tabela 4.13: Izračun MSE v primeru izbranih spremenljivk.

spremenljivka	$mse(\overline{y})$	$mse(\overline{y}_w)$	razlika $(mse(\overline{y}_w) - mse(\overline{y})$
ca1	0,932	1,049386	0,117
ej10	1,912	2,141357	0,229
al03	0,883	0,997693	0,115
doh_gosp	2,064	2,27065	0,207

Vir: EHIS 2007.

Glede na to, da se je srednje kvadratna napaka v vseh prikazanih primerih povečala, uteževanje ni imelo pomembne koristi.

4.6 Vpliv načina anketiranja na neodgovor spremenljivke

V zadnjih desetletjih je prišlo do razmaha različnih načinov anketiranja. Raziskovalci lahko izbirajo med načini anketiranja, pri katerih je anketiranje izvedeno s pomočjo anketarja, kot sta osebno in telefonsko anketiranje, ter metodami samoanketiranja, kot so samoanketiranje z vprašalnikom na papirju, anketiranje z anketami po pošti ter spletno anketiranje. Izbor načina anketiranja je en izmed pomembnejših izzivov v anketnem raziskovanju. Raziskovalci se namreč pogosto sprašujejo, kateri način anketiranja ali skupek različnih načinov anketiranja je najprimerneje uporabiti v dani situaciji (De Leeuw 2005). Načini anketiranja se med seboj razlikujejo predvsem v stopnji stika med anketirancem in raziskovalcem ali anketarjem ter v količini in vzrokih za pojav neodgovora, kakovosti zbranih podatkov in v stroških (Hunt-White 2006). Raziskave vpliva načina anketiranja se nanašajo predvsem na to, kje, kdaj in na kakšen način se raziskava izvaja (Fendrich in Johnson 2001). Ob primerjavi različnih načinov anketiranja se največje razlike pojavijo med posameznimi načini, kadar gre za občutljivejša vprašanja. Pri občutljivejših vprašanjih anketiranje s pomočjo anketarja namreč povzroča večje deleže družbeno zaželenih odgovorov in več nedoslednih

odgovorov pa tudi bolj natančne odgovore v primeru odprtih vprašanj, kar se izrazito pokaže v primeru osebnega in tudi telefonskega anketiranja (De Leeuw 2005). Vendar zakaj način anketiranja sploh vpliva na odgovore anketirancev? Če so anketiranci prepričani glede posameznega odgovora na neko vprašanje, bi vendar morali na vprašanje pravilno odgovoriti ne glede na način zbiranja podatkov.

A raziskave kažejo, da zasebnost v procesu odgovarjanja pomembno vpliva na rezultate občutljivih anketnih vprašanj, kot so vprašanja o nedovoljenih in tveganih vedenjih, npr. uporaba drog, ter pri občutljivih psiholoških temah. Anketiranci namreč na taka vprašanja lažje odgovarjajo, kadar jim odgovorov ni treba razkriti anketarju. Družbeno zaželeno odgovarjanje, ki lahko za raziskovalce predstavlja velik problem v primeru osebnega anketiranja, se ob uporabi samoanketiranja pomembno zmanjšuje, kar pa verjetno nima vpliva, kadar gre za manj občutljiva vsebinska področja (Aquilino 1998). Primerjalne raziskave namreč ne ugotavljajo vpliva načina anketiranja, kadar gre za splošna vprašanja s področja zdravja (npr. Groves in drugi; Aneshensel in drugi v Aquilino 1998). Na razsežnosti kakovosti podatkov, kot so natančnost, točnost in neodgovor spremenljivke, vplivajo tri psihološke spremenljivke, in sicer zasebnost, zakonitost in kognitivni napor. Na te psihološke spremenljivke pa pomembno vpliva način anketiranja, torej dejstvo, ali so podatki zbrani z osebnim anketiranjem, samoanketiranjem, spletnim ali telefonskim anketiranjem (Essig in Winter 2009).

Poleg vpliva psiholoških spremenljivk na kakovost zbranih podatkov ne smemo prezreti tudi vizualnega oziroma organizacijskega učinka samega vprašalnika, ki je še posebej izrazit v primeru anketnega vprašalnika, namenjenega samoanketiranju. Pomemben vpliv na pojav neodgovorov in nekonsistentnih odgovorov ima tudi organizacija samega vprašalnika, kar je še posebej izrazito v primeru vprašalnika na papirju. Jenkins in Dillman (1997) ugotavljata pomembnost zastavljanja samo enega vprašanja

naenkrat. Slednje je namreč predvsem izrazito v primeru samoanketiranja z vprašalnikom na papirju, kjer je predvsem zaradi zmanjšanja stroškov izvedbe anketne raziskave to pogosto prezrto, saj se ponavadi teži k čim manjši dolžini vprašalnika v smislu obsega strani, zaradi česar je na eno stran stisnjenih kar se da veliko število vprašanj.

Groves in Peytcheva (2008) velik vpliv na pojav neodgovorov pri anketiranju s pomočjo anketarja pripisujeta načinu rekrutacije respondentov, pri samoanketiranju pa predvsem skladnosti navodil za anketiranje s prošnjo za sodelovanje v anketi.

Essig in Winter (2009) v raziskavi med splošno populacijo ugotavljata, da so deleži neodgovorov na bolj občutljiva anketna vprašanja višji pri računalniško podprtem osebnem anketiranju, nižji pa pri samoanketiranju z vprašalnikom na papirju, kjer lahko anketiranci v zasebnosti svojega doma samostojno odgovorijo na občutljivejša anketna vprašanja, kot so npr. vprašanja o dohodku. Jäckle in drugi (2006) pa ugotavljajo, da so v splošnem stopnje neodgovora spremenljivke višje pri osebnem kot pri telefonskem anketiranju. Najvišje deleže neodgovorov ugotavljajo pri vprašanjih o dohodku, pri čemer so stopnje neodgovorov višje pri osebnem anketiranju. Po drugi strani pa so deleži odgovorov »ne vem« večji pri telefonskem kakor pri osebnem anketiranju.

Način anketiranja v anketah, povezanih z zdravjem, vpliva tudi na pojav opustitve odgovarjanja v neki točki vprašalnika, kar je ena izmed oblik neodgovora spremenljivke. Višji deleži opustitve odgovarjanja v neki točki se namreč pojavljajo pri poštnih anketah oziroma samoanketiranju, medtem ko so pri telefonskih anketah in osebnem anketiranju na domu respondenta ti deleži bistveno nižji (Gilbert in drugi 1992).

Raziskovalci s področja zdravja in vedenj, povezanih z zdravjem, se pogosto zanašajo na samoporočane podatke, čeprav se zavedajo

morebitne netočnosti in nepopolnosti zbranih podatkov. Motivacijski dejavniki za sodelovanje in odgovarjanje na posamezna anketna vprašanja pa lahko vodijo k različnim odgovorom znotraj različnih načinov anketiranja, saj se lahko pristranskost odgovora pojavi npr. zaradi posameznikove želje po predstavitvi samega sebe na družbeno zaželen način (Newman in drugi 2002). Stopnje neodgovora spremenljivke so pri vprašanjih, povezanih s tveganimi vedenji, višje pri zbiranju podatkov s samoanketiranjem, kot pa pri zbiranju podatkov s pomočjo telefonskega anketiranja (Aquilino 1994).

Brener in drugi (2006) v svoji raziskavi vpliva načina anketiranja na poročanje podatkov o tveganih vedenjih med otroki in mladimi v šolskem obdobju ugotavljajo, da otroci in mladostniki, ki so anketo izpolnjevali s samoanketiranjem preko spleta, pogosteje poročajo različna tvegana vedenja kot otroci in mladostniki, ki so anketo izpolnjevali s samoanketiranjem z vprašalnikom na papirju. Ugotavljajo tudi pomemben vpliv načina anketiranja glede na občutljivost posameznih anketnih vprašanj, in sicer na poročanje podatkov o poškodbah, pitju alkohola, uporabi prepovedanih drog, kajenju in nadzoru nad telesno težo. Način anketiranja torej pri otrocih in mladostnikih pomembno vpliva na poročanje odgovorov o bolj občutljivih vedenjih.

Vpliv načina anketiranja na pojav neodgovora spremenljivke v »Anketi o zdravju in zdravstvenem varstvu«

V nadaljevanju bomo v primeru »Ankete o zdravju in zdravstvenem varstvu« preverjali vpliv ocenjene občutljivosti posameznih izbranih anketnih vprašanj (občutljivejša vprašanja v primerjavi z manj občutljivimi) in načina anketiranja (samoanketiranje v primerjavi z anketiranjem s pomočjo anketarja z vprašalnikom na papirju), na pojav neodgovora spremenljivke.

Spremenljivko *način anketiranja* smo izračunali tako, da smo izbrana anketna vprašanja najprej razdelili v dva sklopa glede na način na katerega so bili podatki zbrani, in sicer smo posameznim izbranim ključnim vprašanjem pripisali vrednost 1, v primeru kadar je šlo za anketiranje s pomočjo anketarja z vprašalnikom na papirju in 2 kadar je šlo za samoanketiranje z vprašalnikom na papirju. Vprašanja o kajenju, alkoholu in drogah so namreč ocenjena kot izredno občutljiva, zato je bilo treba pri izvedbi ankete upoštevati priporočilo Eurostata, da se pri teh vprašanjih uporabi zbiranje podatkov s samoanketiranjem z vprašalnikom na papirju. Anketarji so imeli navodilo, da vprašalnik na papirju s sklopom vprašanj o »Kajenju, alkoholu in drogah«, anketiranci izpolnijo sami. Anketar je pred pričetkom anketiranja v desni zgornji kot ankete vpisal šifro izbrane osebe, ter poleg vprašalnika za samoanketiranje anketirancu izročil tudi kuverto, v katero je nato anketiranec vložil izpolnjen vprašalnik (Lavtar in drugi 2007).

Vsaki izbrani spremenljivki smo nato pripisali vrednost 0, če je oseba odgovorila na izbrano vprašanje in 1, če ni (neodgovor spremenljivke). Nato smo neodgovore spremenljivke sešteli v skupno število neodgovorov glede na način anketiranja (neodgovor spremenljivke v primeru samoanketiranja proti neodgovoru spremenljivke v primeru podatkov zbranih s pomočjo anketarja). Spodnji tabeli prikazujeta število neodgovorov spremenljivke glede na način anketiranja.

Tabela 4.14: Število neodgovorov spremenljivke v primeru izbranih vprašanj in odgovorov zbranih s pomočjo anketarja.

	Pojav neodgovora spremenljivke	**n**	**%**
0	Odgovor spremenljivke	1465	69,2
1	Neodgovor spremenljivke	104	4,9
2		454	21,4
3		78	3,7
5		1	0,0
9		1	0,0
10		1	0,0
11		7	0,3
12		7	0,3
Skupaj		2118	100,0

Vir: EHIS 2007.

Opazimo lahko, da se v primeru izbranih spremenljivk višja stopnja neodgovora spremenljivke pojavlja pri vprašanjih zbranih s samoanketiranjem.

Tabela 4.15: Število neodgovorov spremenljivke v primeru izbranih vprašanj in odgovorov zbranih s samoanketiranjem.

Št. neodgovorov	**Pojav neodgovora spremenljivke**	**n**	**%**
0	Odgovor spremenljivke	1118	52,8
1	Neodgovor spremenljivke	776	36,6
2		27	1,3
3		45	2,1
4		79	3,7
5		52	2,5
6		5	0,2
7		10	0,5
9		2	0,1
10		4	0,2
Skupaj		2118	100,0

Vir: EHIS 2007.

Na osnovi zgoraj navedenega števila neodgovorov smo nato za vsak način anketiranja ločeno izračunali spremenljivki *pojav neodgovora*

spremenljivke, kjer 0 pomeni odgovor spremenljivke, 1 pa neodgovor spremenljivke, in sicer:

- neodgovor spremenljivke v primeru vprašanj zbranih s pomočjo anketarja in
- neodgovor spremenljivke v primeru vprašanj zbranih s samoanketiranjem (oboje z vprašalnikom na papirju).

Podatke smo nato zapisali v dva ločena seta podatkov, ki sta oba vsebovala naslednje spremenljivke:

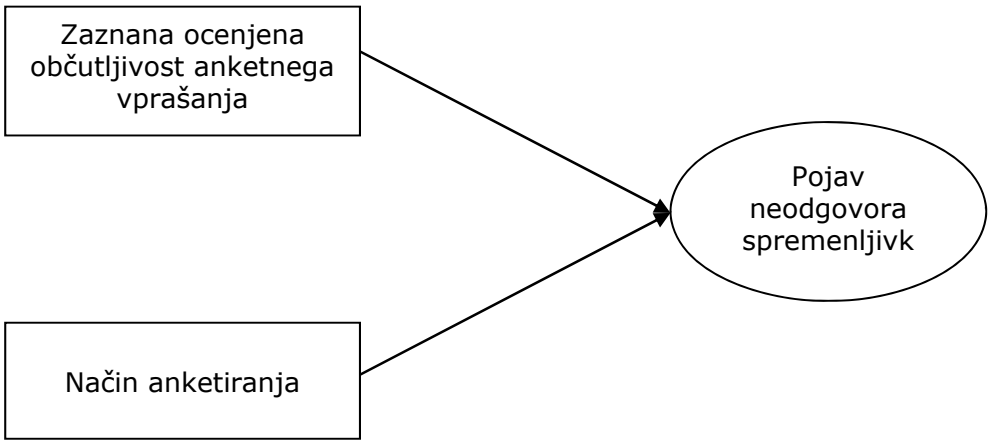
- spol,
- starostne skupine,
- dosežena stopnja izobrazbe,
- formalni zakonski stan,
- trenutni zaposlitveni status,
- stopnja urbanizacije in
- način anketiranja (samoanketiranje, v primerjavi z anketiranjem s pomočjo anketarja z vprašalnikom na papirju) in
- ocenjena občutljivost vprašanj (1 – občutljivo vprašanje, 2 – neobčutljivo vprašanje).

Spremenljivko *zaznana ocenjena občutljivost anketnega vprašanja* smo izračunali tako, da smo najprej izbranim anketnim vprašanjem, ki jih tudi v literaturi označujejo kot občutljiva (vprašanja o pitju alkoholnih pijač, uživanju marihuane in skupnem dohodku gospodinjstva) pripisali oznako *občutljivo anketno vprašanje*, nato pa za vsakega posameznika izračunali skupen seštevek neodgovora spremenljivke za izbrana ocenjena občutljiva anketna vprašanja in mu v primeru, da posameznik na vsaj eno izmed izbranih vprašanj ni odgovoril, pripisali vrednost 1, ki pomeni zaznavanje anketnega vprašanja ocenjenega kot občutljivega ali pa vrednost 2, kadar je anketiranec odgovoril na vsa izbrana anketna vprašanja ocenjena kot

občutljiva (ocena nezaznane občutljivosti anketnega vprašanja kot občutljivega)[12].

Za potrebe nadaljnje analize smo ločena seta podatkov združili v enega. Vpliv zaznane ocenjene občutljivosti anketnih vprašanj in načina anketiranja na pojav neodgovora spremenljivke, bomo v nadaljevanju preverjali s pomočjo logistične regresije.

Slika 4.12: Raziskovalni model vpliva zaznane ocenjene občutljivosti anketnega vprašanja in načina anketiranja na pojav neodgovora spremenljivke v EHIS.



Ker je en izmed zahtevanih pogojev za izvedbo logistične regresije neobstoj kolinearnosti med opazovanimi pojasnjevalnimi spremenljivkami, smo najprej izvedli test multikolinearnosti. Glede na faktorje VIF, lastne vrednosti in z majhnimi lastnimi vrednostmi povezane nizke deleže variance ugotavljamo, da odvisnost med opazovanimi pojasnjevalnimi spremenljivkami ne obstaja, zato lahko nadaljujemo z analizo s pomočjo logistične regresije.

V model logistične regresije smo vključili spremenljivki *zaznana ocenjena občutljivost anketnega vprašanja* in *načina anketiranja,* za kateri na podlagi preteklih raziskav predvidevamo, da bi lahko imeli vpliv na pojav neodgovora spremenljivke. Model je kot celota statistično značilen

[12] Predpostavljamo, da anketiranec, ki ne odgovori na določeno anketno vprašanje, čeprav bi moral, le-tega ne stori zato, ker vsebino anketnega vprašanja zaznava kot občutljivo.

(p<0,001), Hosmer-Lemeshow test pa kaže, da se od podatkov ne razlikuje statistično značilno (p=0,147). Z modelom smo uspeli pojasniti 38,3 %[13] variabilnosti odvisne spremenljivke. Rezultati logistične regresije so prikazani v spodnji tabeli.

Tabela 4.16: Model logistične regresije (model 2).

	Ocena parametra (β)	Standardna napaka	Wald	SP	Stat. značilnost	Razmerje obetov (exp(β))	95% IZ[14] za razmerje obetov	
							Spodnja meja	Zgornja meja
Način anketiranja								
samoanketiranje	0,977	0,078	158,669	1	0,000	2,656	2,281	3,092
Ocenjena občutljivost anketnega vprašanja								
ocenjena občutljivost	2,611	0,084	974,362	1	0,000	13,610	11,552	16,034
Konstanta	-2,512	0,086	862,041	1	0,000	0,081		

Na podlagi statistične značilnosti Waldove statistike ugotavljamo, da tako način anketiranja, kakor tudi zaznana ocenjena občutljivost anketnega vprašanja s strani anketiranca vplivata na pojav neodgovora spremenljivke. Zbiranje podatkov s pomočjo samoanketiranja ima namreč 2,7-krat višji obete za pojav neodgovora spremenljivke, kot zbiranje podatkov s pomočjo anketarja (95 % IZ je od 2,28 do 3,10). Anketna vprašanja za katera se predvideva, ki jih anketiranci zaznavajo kot občutljiva, pa imajo 13,6-krat višje obete za pojav neodgovora spremenljivke, kot vprašanja za katera se ocenjuje, da jih anketiranci ne zaznavajo kot občutljiva (95 % IZ je od 11,55 do 16,03).

Analiza neodgovora spremenljivke, ki jo je opravil Kveder (2005) ugotavlja podobno. Višji deleži neodgovora spremenljivke se bolj verjetno pojavljajo v primeru občutljivejših vprašanj o spolnih in drugih vedenjih ter vprašanjih, ki s strani anketiranca zahtevajo, da se spominja določene podatke o dogodkih, ki so se zgodili v preteklosti.

[13] Nagelkerke R^2
[14] Interval zaupanja

5 RAZPRAVA IN ZAKLJUČEK

Padajoče stopnje neodgovora enote v anketnih raziskovanjih lahko ob sočasnem pojavu neodgovora spremenljivke pomembno vplivajo na kakovost izvedenih anket in izračunane ocene populacijskih parametrov. V magistrski nalogi smo celostno obravnavali različne dejavnike, ki vplivajo na pojav neodgovora spremenljivke, pri čemer je bil naše osnovno izhodišče posameznik, njegove značilnosti in nekateri zunanji dejavniki, ki vplivajo na pojav neodgovora spremenljivke.

Značilnosti anketirancev pomembno vplivajo na pojav neodgovora spremenljivke. Tako se višji deleži neodgovora spremenljivke pojavljajo pri ženskah, starejših anketirancih, srednje in nižje izobraženih, neporočenih in posameznikih z nižjimi osebnimi dohodki. Rezultati analiz so le delno skladni z ugotovitvami predhodnih raziskav, saj ugotavljamo, da imata izmed vseh analiziranih socio-demografskih dejavnikov pomemben vpliv na pojav neodgovora spremenljivke le spol in izobrazba. Ženske imajo namreč višje obete za pojav neodgovora spremenljivke, kot moški. Višje izobraženi posamezniki pa imajo višje obete za pojav neodgovora spremenljivke kot nižje izobraženi, kar se ne sklada s predhodnimi ugotovitvami. Razlog za to morda izhaja iz različnega načina zaznavanja občutljivosti posameznih anketnih vprašanj pri bolj izobraženih posameznikih. Morda višje izobraženi posamezniki nekatera izbrana ključna vprašanja zaznavajo kot bolj občutljiva in odgovorov nanje enostavno ne želijo razkriti tretjim osebam. Morda pa so višje izobraženi posamezniki v »Anketi o zdravju in zdravstvenem varstvu« ženskega spola in moramo opazovati tudi sočasen vpliv izobrazbe in spola na neodgovor spremenljivke.

Glede na rezultate logistične regresije, kjer smo obravnavali vpliv socio-demografskih značilnosti posameznikov na pojav neodgovora spremenljivke, hipoteze o vplivu starosti na pojav neodgovora ne moremo

potrditi. Predvidevamo, da so dobljeni rezultati pogojeni s tematiko same anketne raziskave in posameznih izbranih anketnih vprašanj, kjer neodgovor na posamezna anketna vprašanja ni nujno pogojen s starostjo anketiranca (npr. anketna vprašanja o pitju alkohola, kjer so bile stopnje neodgovora na anketno vprašanje v Anketi o zdravju in zdravstvenem varstvu najvišje). Ob opazovanju samostojnega vpliva starosti na pojav neodgovora spremenljivke sicer ugotavljamo, da starost vpliva na pojav neodgovora spremenljivke, saj se z nižanjem starosti višajo obeti za pojav neodgovora spremenljivke, vendar z modelom pojasnimo le 0,4 % variabilnosti pojava neodgovora, kar pomeni, da starost anketiranca v tem primeru ne predstavlja pomembnega dejavnika.

Tudi Kveder (2005) ugotavlja podobno kot je pokazala naša analiza, in sicer v analizi treh različnih anketnih raziskav ugotavlja, da se večji deleži neodgovora spremenljivke pojavljajo pri ženskah, vendar se vpliv spola povezuje tudi s tematiko anketne raziskave in zanimanjem za odgovarjanje na posamezna anketna vprašanja. Njegova analiza vpliva izobrazbe na neodgovor spremenljivke je pokazala ravno nasproten vpliv, kot ga ugotavljamo v »Anketi o zdravju in zdravstvenem varstvu«. Ugotavlja namreč, da se stopnje neodgovora spremenljivke višajo z nižanjem izobrazbe posameznika. Njegova analiza je pokazala tudi vpliv starosti na neodgovor spremenljivke, in sicer se višji deleži neodgovora spremenljivke pojavljajo pri starejših anketirancih.

Ena pomembnejših spremenljivk, s pomočjo katerih se merijo neenakosti v zdravju, je tudi dohodek gospodinjstva, saj posledično kaže na socialno-ekonomski položaj posameznika v družbi. Opažamo, da tudi pri združeni spremenljivki o skupnem dohodku gospodinjstva, kjer sta združena odprto vprašanje in vprašanje o posameznih dohodkovnih razredih, delež neodgovorov še vedno predstavlja 25,8 % vseh odgovorov, zato je nujno, da se preučijo načini, s pomočjo katerih bi lahko v prihodnjih anketnih raziskovanjih o zdravju znižali deleže neodgovora spremenljivke na

vprašanje o dohodku, da bomo lahko zagotovili kakovostne mere, s katerimi bomo lahko opazovali neenakosti v zdravju. Po drugi strani pa očitno mladi vprašanj, s katerimi merimo njihov socialno-ekonomski položaj in socialno-ekonomski položaj njihove družine v družbi, ne zaznavajo kot občutljivih, saj so deleži neodgovora spremenljivke zanemarljivi, vendar pa so te mere pri mladih pogosto vprašljive kakovosti. Mladi namreč v anketnih raziskovanjih pogosto poročajo nekonsistentno, kar je razvidno tudi v primeru nekaterih anketnih vprašanj v anketi »Z zdravjem povezano vedenje v šolskem obdobju«. Med mladimi se sicer najvišji deleži neodgovorov spremenljivke pojavljajo pri različnih tveganih vedenjih, kar potrjujejo tudi naše analize, kjer opažamo, da se največja deleža neodgovorov pojavita pri vprašanjih o spolnih odnosih.

Tudi umeščenost anketnih vprašanj v vprašalniku in njihova tematika pomembno vplivata na stopnje neodgovorov v raziskavah o tveganih vedenjih med mladimi. Analize namreč kažejo, da imajo vprašanja, ki so predhodna vprašanjem o spolnih vedenjih, in vprašanja, ki tem vprašanjem sledijo, nižje stopnje neodgovora spremenljivke kot vprašanja o spolnih vedenjih (Saewyc in drugi 2004), kar potrjujejo tudi naše analize.

Neodgovori spremenljivke se pogosteje pojavljajo takrat, ko gre za vprašanja osebne narave, ki jih anketiranec zaznava kot občutljiva. To so npr. različna družbeno nezaželena vedenja, vprašanja o dohodku ipd. V okviru anket o zdravju pa gre za vprašanja o spolnem vedenju, pitju alkoholnih pijač, kajenju, uživanju prepovedanih drog, pa tudi za vprašanja o različnih boleznih in bolezenskih stanjih, ki jih anketiranec enostavno ne želi razkriti tretjim osebam. Podatki o diagnozah bolezni so tudi v ZVOP označeni kot občutljivi osebni podatki, s katerimi je treba ravnati še posebej pazljivo.

Tudi rezultati naše analize kažejo podobno. Opazimo lahko namreč, da se največji delež manjkajočih odgovorov pojavlja v primeru vprašanja o tveganih oziroma družbeno nezaželenih vedenjih, torej v primeru popitih šestih mericah alkoholnih pijač ob eni priložnosti v zadnjih 12 mesecih, kjer neodgovori predstavljajo skoraj polovico vseh odgovorov. Visoki deleži neodgovora spremenljivke so tudi v primerih pitja alkoholnih pijač v posameznih dnevih tedna (do 39,4 %). Medtem ko vprašanja o posameznih dohodkih gospodinjstva z vidika neodgovorov niso problematična, saj neodgovori v nobenem primeru ne presegajo 1 %, pa vprašanja o skupnem dohodku gospodinjstva očitno pri anketirancih zbujajo nepripravljenost na odgovarjanje. Deleži neodgovorov so namreč visoki, saj tudi pri združeni spremenljivki, kjer sta združena odprto vprašanje in vprašanje o posameznih dohodkovnih razredih, delež neodgovora spremenljivke še vedno predstavlja četrtino vseh odgovorov. Po drugi strani pa neodgovori spremenljivke v primeru ostalih vprašanj, npr. o splošnem zdravstvenem stanju, posameznih izbranih boleznih in bolezenskih stanjih, zdravilih ipd., ne presegajo 1 %. Zgornje ugotovitve lahko podkrepimo še z rezultati logistične regresije, kjer ugotavljamo, da imajo anketna vprašanja, katera ocenjujemo, da jih posamezniki zaznavajo kot občutljiva, višje obete za pojav neodgovora spremenljivke.

V pregledu literature smo ugotovili, da tudi način anketiranja pomembno vpliva na neodgovor spremenljivke, saj se večji deleži neodgovora spremenljivke pojavljajo pri samoanketiranju, vpletenost anketarja v proces anketiranja pa ima ob zagotavljanju ustreznega izobraževanja anketarjev pozitiven vpliv na deleže neodgovora spremenljivke, ki se kaže predvsem pri zapletenih anketnih vprašanjih, ki neukemu anketirancu pri samoanketiranju povzročajo mnogo težav, raziskovalcem pa probleme pri analizi zbranih podatkov, zaradi visoke stopnje neodgovora spremenljivke (slednje velja predvsem za zahtevna vprašanja, kjer je zahtevana sočasna uporaba preskokov). Z rezultati logistične regresije, kjer smo obravnavali vpliv ocenjene občutljivosti anketnega vprašanja in načina anketiranja na

pojav neodgovora na anketno vprašanje, lahko potrdimo hipotezo o vplivu načina anketiranja na pojav neodgovora spremenljivke. Ugotovili smo namreč, da so obeti za pojav neodgovora spremenljivke višji pri samoanketiranju z vprašalnikom na papirju kot pri osebnem anketiranju z vprašalnikom na papirju. Tudi analiza neodgovora spremenljivke, ki jo je opravil Kveder (2005) ugotavlja podobno. Višji deleži neodgovora spremenljivke se bolj verjetno pojavljajo v primeru občutljivejših vprašanj o spolnih in drugih vedenjih ter vprašanjih, ki s strani anketiranca zahtevajo, da se spominja določene podatke o dogodkih, ki so se zgodili v preteklosti.

Ugotovljena splošna stopnja neodgovora spremenljivke v obeh anketah je okoli 1 % do 2 %. Najvišja stopnja neodgovora spremenljivke se v »Anketi o zdravju in zdravstvenem varstvu« pojavi pri spremenljivki o popitih šestih mericah alkoholnih pijač ob eni priložnosti v zadnjih 12 mesecih, kjer neodgovor spremenljivke predstavlja skoraj polovico vseh odgovorov (46,2 %). V anketi »Z zdravjem povezano vedenje v šolskem obdobju« pa se najvišja stopnja neodgovora spremenljivke, pojavi pri spremenljivki o starosti ob prvem spolnem odnosu, kjer znaša 3,2 %. Čeprav je splošna stopnja neodgovora spremenljivke relativno nizka, pa lahko tudi pojav relativno nizkih stopenj neodgovora spremenljivke (1 % do 2 %) pomembno vpliva na kakovost dobljenih rezultatov, zato je pomembno, da tudi v takšnih primerih probleme zaradi pojava neodgovora spremenljivke, rešujemo z metodami vstavljanja podatkov (imputacijo).

Reševanje pojava neodgovora spremenljivke zagotovo ni enostavno, začne pa se že v fazah načrtovanja ankete in priprave anketnega vprašalnika, kjer je treba uporabiti tudi ustrezne oblike vprašanj, ki so primernejše za specifične populacijske skupine. Anketarji morajo biti zadostno izobraženi, ustrezna in dovolj natančna navodila pa so potrebna tudi pri samoanketiranju, kjer pa moramo paziti tudi na njihov obseg, saj

lahko obsežna navodila dodatno povzročajo neodgovor spremenljivke. Seveda ne smemo pozabiti tudi na metode vstavljanja in uteževanja, s katerimi vpliv neodgovora spremenljivke rešujemo po opravljeni terenski fazi ankete in ki lahko pomembno vplivajo na kakovost dobljenih rezultatov.

6 LITERATURA

1. Alvik, Astrid, Haldorsen, Tor in Lindermann, Rolf. 2005. Consistency od Reported Alcohol Use by Pregnent Women: Anonymous Versus Questionnaires With Item Nonresponse Differences. *Alcoholism: Clinical and Experimental Research* 29 (8): 1444-1449.
2. Andridge, Rebecca R. in Little, Roderick J. 2009. The Use of Sample Weights in the Hot Deck Imputation. *Jurnal of Official Statistics* 25 (1): 21-36.
3. Aquilino, William S. 1994. Interview mode effects in Surveys of drug and alcohol use. The Public Opinion Quarterly 58 (2): 210-240.
4. Aquilino, William S. 1998. Effects of Interview Mode on Measuring Depression in Younger Adults. *Journal of Official Statistics* 14 (1): 15-29.
5. Barnett, Julie. 1998. Sensitive questions and response effects: an evaluation. Journal of Managerial Psychology 13 (1/2): 63-76.
6. Baruch, Yebuda in Holton, Brooks C. 2008. Survey response rate levels and trends in organizational research. *Human Relations* 61 (8): 1139-1160.
7. Beatty, Paul in Herrmann, Douglas. 2002. To Answer or Not to Answer: Decision Processes Related to Survey Item Nonresponse. V *Survey Nonresponse*, ur. Groves, Robert M., Dillman, Don A., Eltinge, John L. in Little, Roderick J.A., 3-26. New York: John Wiley and Sons Inc.

8. Beebe, Timothy J., Jenkins, Sarah M., Anderson, Kari J. in Davern, Michael E. 2008. Survey-Related Experimental and Attitudinal Correlates of Future Health Survey Participation: Results of a Statewide Survey. *Mayo Clinic Proceedings* 83 (12): 1358-1363.
9. Bethlehem, J. G. in Kersten, H. M. P. 1985. On the Treatment of Nonresponse in Sample Surveys. Journal of Official Statistics 1 (3): 287-300.
10. Borgers, Natacha in Hox, Joop. 2001. Item Nonresponse in Questionnaire Research With Children. *Journal of Official Statistics* 17 (2): 321-335.
11. Borgers, Natacha, de Leeuw, Edith in Hox, Joop. 2000. Children as Respondents in Survey Research: Cognitive Development and Response Quality*. Bulletin de Méthodologie Sociologique* 66: 60-75.
12. Bradburn, Norman M., Sudman, Seymour, Blair, Ed in Stocking, Carol. 1978. Question Threat and Response Bias. *Public Opinion Quarterly* 42 (2): 221-234.
13. Brener, Nancy D., Eaton, Danice K., Kann, Laura, Grunbaum, Jo Anne, Gross, Lori, A., Kyle, Tonja M. in Ross, James G. 2006. The association of survey setting and mode with self-reported health risk behaviours among high school students. *Public Opinion Quarterly* 70 (3): 354-374.
14. Buzeti, Tanja, Gabrijelčič Blenkuš, Mojca, Gruntar Činč, Mojca, Ivanuša, Marijan, Pečar, Janja, Tomšič, Sonja, Truden Dobrin, Polonca in Vrabič Kek, Brigita, ur. 2011. *Neenakosti v zdravju v Sloveniji*. Ljubjana: Inštitut za varovanje zdravja Republike Slovenije.
15. Cohen, G. In Duffy, J.C. 2002. Are Nonrespondents to Health Surveys Less Healthy Than Respondents? *Journal of Official Statistics* 18 (1): 13-23.
16. Davern, Michael, Rodin Holly, Beebe, Timothy J. in Thiede Call, Kathleen. 2005. The Effect of Income Question Design in Health Surveys on Family Income, Poverty and Eligibility Estimates. *Health Services Research* 40 (5): 1534-1552.

17. De Leeuw, Edith D. 1992. Data Quality in Mail, Telephone and Face to Face Surveys. Amsterdam: TT-Publikaties Amsterdam.
18. De Leeuw, Edith D., Hox, Joop J. in Dillman, Don D., ur. 2008. *International Handbook of Survey Methodology*. New York: Taylor and Francis Group.
19. De Leeuw, Edith in de Heer, Wim. 2002. Trends in Household Survey Nonresponse: A Longitudinal and International Comparison. V *Survey Nonresponse*, ur. Groves, Robert M., Dillman, Don A., Eltinge, John L. in Little, Roderick J.A., 3-26. New York: John Wiley and Sons Inc.
20. De Leeuw, Edith, Hox, Joop in Huisman, Mark. 2003. Prevention and treatment of Item Nonresponse. *Journal of Official Statistics* 19 (2): 153-176.
21. De Leeuw, Edith. 2001. Reducing Missing Data in Surveys: An Overview of Methods. *Quality & Quantity* 35: 147-160.
22. De Leeuw, Edith. 2005. To Mix or Not to Mix Data Collection Modes in Surveys. *Journal of Official Statistics* 21 (2): 233-255.
23. De Leeuw, Edith. 2011. Improving Data Quality when Surveying Children and Adolescents: Cognitive and Social Development and its Role in Questionnaire Construction and Pretesting. *Report prepared for the Annual Meeting of the Academy of Finland: Research Programs Public Health Challenges and Health and Welfare of Children and Young People*. Utrecht University. Dostopno prek: http://www.aka.fi/Tiedostot/Tiedostot/LAPSET/Presentations%20of%20the%20annual%20seminar%2010-12%20May%202011/Surveying%20Children%20and%20adolescents_de%20Leeuw.pdf (29. avgust 2011).
24. Dengler, Regina, Roberts, Heather in Rushton, Lesley. 1997. Lifestyle surveys – the complete answer? *Journal of Epidemiology and Community Health* 51: 46-51.
25. Deville, Jean- Claude in Särndal, Carl E. 1994. Variance Estimation for the Regression Imputed Horvitz-Thompson Estimator. *Journal of Official Statistics* 10 (4): 381-394.

26. Dillman, Don A., Eltinge, John L., Groves, Robert M. in Little, Roderick J.A. 2002. Survey Nonresponse in Design Data Collection and Analysis. V *Survey Nonresponse*, ur. Groves, Robert M., Dillman, Don A., Eltinge, John L. in Little, Roderick J.A., 3-26. New York: John Wiley and Sons Inc.
27. Durrant, Gabriele B. 2005. Imputation Methods for Handling Item-Nonresponse in the Social Sciences: A Methodological Review. ESRC National Centre for Research Methods and Southampton Statistical Sciences Research Institute, University of Southampton. *NCRM Methods Review Papers*. Dostopno prek: http://www.ncrm.ac.uk/research/outputs/publications/methodsreview/MethodsReviewPaperNCRM-002.pdf (5. maj 2011).
28. Ekholm, Ola, Gundgaard, Jens, Rasmussen, Niels K.R. in Holme Hansen, Ebba. 2010. The effect of health, socio-economic position and mode of data collection on non-response in health interview surveys. *Scandinavian Journal of Public Health* 38: 699-706.
29. Elliott, Marc N, Edwards, Carol, Angeles, January, Hambarsoomians, Katrin in Hays, Ron D. 2005. Petterns of Unit and Item Nonresponse in the CAHPS Hospital Survey. *Health Services Research* 40 (6): 2096-2119.
30. Essig, Lothar in Winter, Joachim K. 2009. Item Non-Response to Financial Questions in Household Surveys: An Experimental Study of Interviewer and Mode Effects. *Fiscal Studies* 30 (3/4): 367-390.
31. Fendrich, Michael in Johnson, Timothy P. 2001. Examining prevalence differences in three national surveys of youth: Impact of consent procedures, mode and editing rules. *Journal of Drug Issues* 31 (3): 615-642.
32. Fricker, Scott in Tourangeau, Roger. 2010. Examining the relationship between nonresponse propensity and data quality in two national household surveys. *Public Opinion Quarterly* 74 (5): 934-955.
33. Gribble, James N., Miller, Heater G., Cooley, Philip C., Catania, Joseph A., Pollack, Lance in Turner, Charles F. 2000. The impact of T-ACASI

interviewing on reported drug use among men who have sex with men. *Substance Use and Misuse* 35 (6-8):869-890.

34. Gilbert, Gregg H., Longmate, Jeffrey in Branch, Laurence G. 1992. Factors Influencing the Effectiveness of Mailed Health Surveys. *Public Health Report* 107 (5): 576-584.
35. Goyder, John, Warriner, Keith in Miller, Susan. 2002. Evaluating Socio-economic Status (SES) Bias in Survey Nonresponse. *Journal of Official Statistics* 18 (1): 1-11.
36. Greenlees, John S., Reece, William S. in Zieschang, Kimberly D. 1982. Imputation of Missing Values When the Probability of Response Depends on the Variable Being Imputed. *Journal of the American Statistical Association* 77 (378): 251-261.
37. Griesler, Pamela C., Kandel, Denise B., Schaffran, Christine, Hu, Mei-Chen in Davis, Mark. 2008. Adolescents' inconsistency in self-reported smoking. *Public Opinion Quarterly* 72 (2): 260-290.
38. Groves, Robert M. 2004. *Survey Errors and Survey Costs*. New Jersy: John Wiley and Sons Inc.
39. Groves, Robert M. 2006. Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly* 70 (5): 646-675.
40. Groves, Robert M. in Peytcheva, Emilia. 2008. The impact of nonresponse rates on nonresponse bias. *Public Opinion Quarterly* 72 (2): 167-189.
41. Groves, Robert M., Dillman, Don A., Eltinge, John L. in Little, Roderick J.A. ur. 2002. *Survey Nonresponse*. New York: John Wiley and Sons Inc.
42. Groves, Robert M., Flower, Floyd J., Couper, Mick P., Lepkowski, James M., Singer, Elenor in Tourangeau, Roger. 2004. *Survey Methodology*. New Jersey: John Wiley and Sons.
43. Heeringa, Steven G., Hill, Daniel H. in Howell, David A. 1995. Unfolding Brackets for Reducing Item Nonresponse in Economic Surveys. PSID A national study of socioeconomics and health over

lifetimes and across generations. *Technical Series Paper* 95-01. Survey Research Center – Institute of Social Research. University of Michigan.

44. Helasoja, Ville, Prättälä, Ritva, Dregval, Liudmila, Pudule, Iveta in Kasmel, Anu. 2002. Late response and item nonresponse in the Finbalt Health Monitor Survey. *European Journal of Public Health* 12: 117-123.
45. Hunt-White, Tracy. 2006. *Why do students participate in surveys? The influence of respondent characteristics, institutional characteristics, and survey design features on survey participation*. Doktorska dizertacija. The Catholic University of America.
46. IVZ. 2008. *Anketa o zdravju in zdravstvenem varstvu*. Dostopno prek: http://www.ivz.si/Mp.aspx?ni=46&pi=5&_5_id=347&_5_PageIndex=0&_5_groupId=185&_5_newsCategory=&_5_action=ShowNewsFull&pl=46-5.0. (12. avgust 2011).
47. Jäckle, Annette, Caroline Roberts in Peter Lynn. 2006. Telephone versus Face-to-Face Interviewing: Mode Effects on Data Quality and Likely Causes. Report on Phase II of the ESS-Gallup Mixed Mode Methodology Project. *ISER Working Paper* 2006-41. Colchester: University of Essex.
48. Jenkins, Cleo R. in Dillman, Don A. 1997. Towards A Theory of Self-Administrated Questionnaire Design. V *Survey Measurement and Process Quality*, ur. Lyberg, Lars E., Biemer, Paul, Collins, Martin, De Leeuw, Edith, Dippo, Cathryn, Schwarz, Norbert in Trewin, Dennis. 165-196, New York: Wiley.
49. Jeriček Klanšček, Helena in Zupanič, Tina. 2011. Raziskava o neenakostih v zdravju in vedenjih, povezanih z zdravjem. V *Neenakosti v zdravju in z zdravjem povezanih vedenjih slovenskih mladostnikov*, ur. Jeriček Klanšček, Helena, Roskar, Saška, Koprivnikar, Helena, Pucelj, Vesna, Bajt, Maja in Zupanič, Tina. 15-23, Ljubljana: Inštitut za varovanje zdravja Republike Slovenije.
50. Juster, Thomas F. in Smith, James P. 1997. Improving the Quality of Economic Data: Lessons from the HRS and AHEAD. *Journal of the American Statistical Association* 92 (440): 1268-1278.

51. Juster, Thomas F., Cao, Honggao, Perry, Michael in Couper, Mick. 2006. *The Effect of Unfolding Brackets on the Quality of Wealth Data in HRS*. Michigan Retirement Research Center Research Paper No. WP 2006-113. University of Michigan. Dostopno prek: http://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID1094805_code908624.pdf?abstractid=1094805&mirid=1 (12. avgust 2011).
52. Kalton, Graham in Flores-Cervantes Ismael. 2003. Weighting Methods. *Journal of Official Statistics* 19 (2): 81-97.
53. Kano, Megumi, Franke, Todd, Afifi, Abdelmonen A. In Bourque, Linda B. 2008. Adequacy of Reporting Results of School Surveys and Nonresponse Effects: A review of the Literature and a Case Study. *Educational Researcher* 37 (8): 480-490.
54. Kaplan, Celia Patricua, Hilton, Joan F., Park-Tanjasiri, Sora in Pérez-Stable, Eliseo J. 2001. The Effect of Data Collection Mode on Smoking Attitudes and Behavior in Young African American and Women: Face-to-face Interview Versus Self-Administered Questionnaires. *Evaluation Review* 25 (4): 454-473.
55. Kautter, John, Khatutsky, Galina, Pope, Gregory C., Chromy, James R. In Adler, Gerald S. 2006. Impact of Nonresponse on Medicare Curent Beneficiary Survey Estimates. *Health care financing review* 27 (4): 71-93.
56. Kish, Leslie. 1965. Survey sampling. John Wiley and Sons.
57. Koivusilta, Leena K, Rimpelä, Arja H in Kautiainen, Susanna M. 2006. Health inequality in adolescence. Does stratification occur by familial social background, family affluence, or personal social position? *BMC Public Health* 6: 110.
58. Košmelj, Katarina. 2001a. Osnove logistične regresije (1. del). Zbornik Biotehniške fakultete Univerze v Ljubljani. *Kmetijstvo* 77: 227-238.
59. Košmelj, Katarina. 2001b. Osnove logistične regresije (2. del). Zbornik Biotehniške fakultete Univerze v Ljubljani. *Kmetijstvo* 77: 239-245.
60. Körmendi, Eszter. 1988. The Quality of Income Information in Telephone and Face to Face Surveys. V *Telephone Survey*

Methodology, ur. Groves, Robert M., Biemer, Paul P., Lyberg, Lars E., Massey, James T., Nicholls, William L. in Waksberg, Joseph. 341-356, New Your: John Wiley and Sons.

61. Krosnick, Jon A. 1999. Survey reasearch. *Annual Review of Psychology* 50: 537-567.
62. Krosnick, Jon A. 2002. The Causes of No-Opinion Responses to Attitude Measures in Surveys: They Are Rarely What They Appear to Be. V *Survey Nonresponse*, ur. Groves, Robert M., Dillman, Don A., Eltinge, John L. in Little, Roderick J.A., 3-26. New York: John Wiley and Sons Inc.
63. Krosnick, Jon A., Holbrook, Allyson L., Berent, Matthew K., Carson, Richard T., Henemann, Michael W., Kopp, Raymond J., Mitchell, Robert, Cameron, Presser, Stanley, Ruud, Paul A., Smith, Kerry V., Moody, Wendy R., Green, Melanie C. in Conaway, Michael. 2002. The impact of »no opinion« response options on data quality. Non-attitude reduction or invitation to satisfice? *Public Opinion Quartely* 66: 371-403.
64. Kupek, Emil. 1998. Determinants of item nonresponse in a large national sex survey. Archives of Sexual Behaviour 27 (6): 581-594.
65. Kveder, Andrej. 2000. Data quality issues in the scope of international comparison FFS survey. V *FFS Flagship conference: partnership and fertility – a revolution?* Bruselj: PAU.
66. Kveder, Andrej. 2005. *Multilevel item nonresponse modeling*. Doktorska dizertacija. Univerza v Ljubljani, Fakulteta za družbene vede.
67. Lavtar, Darja, Moravec Berger, Daša, Trdič, Jana in Zupanič, Tina. 2007. *Anketa o zdravju in zdravstvenem varstvu: Navodila za anketarje*. Interno gradivo. Inštitut za varovanje zdravja Republike Slovenije.
68. Lillard, Lee, Smith, James P. in Welch, Finis. 1986. What Do We Really Know About Wages? The Importance of Nonreporting and Census Imputation. *Journal of Political Economy* 94 (3/1): 489-506.

69. Loosveldt, Geert, Pickery, Jan in Billiet, Jaak. 2002. Item Nonresponse as a Predictor of Unit Nonresponse in a Panel Survey. *Journal of Official Statistics* 18 (4): 545-557.
70. McCluskey, Serena. 2011. Increasing response rates to lifestyle surveys: A pragmatic evidence review. *Perspectives in Public Health* 131 (2): 89-94.
71. McMorris, Barbara J., Petrie, Renee S., Catalano, Richard F., Flemnig, Charles B., Haggerty, Kevin P. in Abbot, Robert D. 2009. Use of Web and In-Person Survey Modes to Gather Data From Young Adults on Sex and Drug Use. *Evaluation Review* 33 (2): 138-158.
72. Messmer, Donald J. in Seymour, Daniel T. 1982. The Effects of Branching on Item Nonresponse. *The Public Opinion Quarterly* 46 (2): 270-277.
73. Moore, Jeffrey, C., Stinson, Linda, L., in Welniak, Jr., Edward, J. 1999. Income Reporting in Surveys: Cognitive Issues and Measurement Error. V *Cognition and Survey Research*, ur. M.G. Sirken, D.J. Herrmann, S. Schechter, N. Schwarz, J.M. Tanur, and R. Tourangeau. New York: Wiley.
74. Newman, Jessica Clark, Des Jarlis, Don C., Turner, Charles F., Gribble, Jay, Cooley, Phillip in Paone, Denise. 2002. The Differential Effects of Face-to-Face and Computer Interview Modes. *American Journal of Public Health* 92 (2): 294-297.
75. Olson, Kristen. 2006. Survey participation, nonresponse bias, measurement error bias, and total error. *Public Opinion Quarterly* 70 (5): 737-758.
76. Owens, Linda in Johnson, Timothy P. in O'Rouke, Diane. 2001. Culture and Item Nonresponse in Health Surveys. V *Seventh Conference on Health Survey Research Methods*, ur. Cynamon, Marcie L. in Kulka, Richard A. 69-74, Maryland: Department of Health and Human Services, Centers for Disease Control and Prevention.

77. Pickery, Jan in Loosveldt, Geert. 2001. An Exploration of Question Characteristics that Mediate Interviewer Effects on Item Nonresponse. *Journal of Official Statistics* 17 (3): 337-350.
78. Pike, Gary R. 2008. Using Weighting Adjustments to Compensate for Survey Nonresponse. *Res High Educ* 49: 153-171.
79. Pleis, John R. in Cohen, Robin A. 2008. *Impact of Income Bracketing on Poverty Measures Used in the National Health Interview Survey's Early Release Program: Perliminary Data from the 2007 NHIS*. CDC/National Center for Health Statistics. Dostopno prek: http://www.cdc.gov/nchs/data/nhis/income.pdf (5. maj 2011).
80. Rässler, Susanne in Riphahn, Regina T. 2006. Survey item nonresponse and its treatment. *Allgemines Statistisches Archiv* 90: 217-232.
81. Riphahn, Regina T. in Serflig, Oliver. 2005. Item non-response on income and wealth questions. *Empirical Economics* 30: 521-538.
82. Rubin, Donald B. 2004. *Multiple Imputation for Nonresponse in Surveys*. New Jersey: John Wiley and Sons Inc.
83. Saewyc, Elizabeth M., Bauer, Greta R., Skay, Carol L., Bearinger, Linda H., Resnick, Michael D., Reis, Elizabeth in Murphy, Aileen. 2004. Measuring Sexual Orientation in Adolescent Health Surveys: Evaluation of Eight School-based Surveys. *Journal of Adolescent Health* 35 (4): 345-360.
84. Sakshaug, Joseph W., Yan, Ting in Tourangeau. 2010. Nonresponse error, measurement error, and mode of data collection: tradeoffs in a multi-mode survey of sensitive and non-sensitive items. *Public Opinion Quarterly* 74 (5): 907-933.
85. Särndal, Carl-Erik in Lundström, Sixten. 2005. *Estimation in Surveys with Nonresponse*. Chichester, West Sussex: John Wiley and Sons Ltd.
86. Saunders, Janne A., Morrow-Howell, Nancy, Spiznagel Edward, Doré, Peter, Proctor, Enola K. in Pescarino, Richard. 2006. Imputing Missing Data: A Comparison of Methods for Social Work Researcers. *Social Work Research* 30 (1): 19-31.

87. Schaeffer, Nora Cate. 2000. Asking questions about threatening topics: A selective overview. V *The science of self-report: Implications for research and practice*, ur. Stone, Arthur A., Turkkan, Jaylan S., Bachrach, Christine A., Jobe, Jared B., Kurtzman, Howard S. in Cain, Virginia J., 105–121. Mahwah: Lawrence Erlbaum Associates, Inc. Publishers.
88. Schouten, Barry. 2007. A Selection Strategy for Weighting Variables Under a Not-Missing-at-Random Assumption. *Journal of Official Statistics* 23 (1): 51-68.
89. Schräpler, Jörg-Peter. 2006. Explaining Income Nonresponse – A Case Study by means of the British Household Panel (BHPS). *Quality & Quantity* 40: 1013-1036.
90. Shoemaker, Pamela J., Eichholz, Martin in Skews, Elizabeth A. 2002. Item nonresponse: Distinguishing between don't know and refuse. *International Journal of Public Opinion Research* 14 (2): 193-201.
91. Singer, Eleanor 2006. Nonresponse bias in household surveys. *Public Opinion Quarterly* 70 (5): 637-645.
92. Stocké, Volker. 2006. Attitudes toward Surveya, Attitude Accessibility and the Effect on Respondents' Susceptibility to Nonresponse. *Quality & Quantity* 40: 259-288.
93. Sudman, Seymour, Bradburn, Norman M. in Schwarz, Norbert. 1996. *Thinking about answers: the application of cognitive processes to survey methodology*. San Francisco: Jossey-Bass.
94. The American Association for Public Opinion Research. 2011. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. 7th edition. AAPOR.
95. Tomaskovic-Devey, Donald, Lieter, Jeffrey in Thompson, Shealy. 1995. Item Nonresponse in Organizational Surveys. *Sociological Methodology* 25: 77-110.
96. Tourangeau, Roger in Smith, Tom W. 1996. Asking Sensitive Questions: The Impact of Data Collection Mode, Question Format, and Question Context. *The Public Opinion Quarterly* 60 (2): 275-304.

97. Tourangeau, Roger in Yan, Ting. 2007. Sensitive Questions in Surveys. *Psychological Bulletin* 133 (5): 859-883.
98. Tourangeau, Roger, Rips, Lance, J. in Rasinski, Kenneth. 2000. *The Psychology of Survey Response*. Cambridge: Cambridge University Press.
99. Vannieuwenhuyze, Jorre, Loosveldt, Geert in Molenberghs, Geert. 2010. A method for evaluating mode effects in mixed.mode surveys. *Public Opinion Quarterly* 74 (5): 1027-1045.
100. Vazquez Alvarez, Rosalia, Melenberg, Bertrand in van Soest, Arthur. 2001. *Nonparametric bounds in the presence of item nonresponse, unfolding brackets, and anchoring*. Dostopno prek: http://arno.uvt.nl/show.cgi?fid=4258 (10. avgust 1011).
101. Vehovar, Vasja. 1995. *Nadomestne enote v anketnem raziskovanju*. Doktorska dizertacija. Univerza v Ljubljani, Ekonomska Fakulteta.
102. Vehovar, Vasja. 2007. *Nepopolni podatki v anketah*. Ljubljana. Neobjavljena verzija.
103. Vehovar, Vasja in Zupanič, Tina. 2007. *Weighting in the ESS – Round 2*. Univerza v Ljubljani, Fakulteta za družbene vede. Interno gradivo.
104. Winglee, Marianne, Kalton, Graham in Rust, Keith. 2001. Handling Item Nonresponse in the U.S. Component of the IEA Reading Literacy Study. *Journal of Educational and Behavioral Statistics* 26 (3): 343-359.
105. Winter, Joachim. 2002. *Bracketing effects in categorized survey questions and the measurement of economic quantities*. University of Mannheim. Dostopno prek: http://www.sfb504.uni-mannheim.de/publications/dp02-35.pdf (10. avgust 1011).
106. Yan, Ting in Curtin, Richard. 2010. The Relation Between Unit Nonresponse and Item Nonresponse: A Response Continuum Perspective. *International Journal of Public Opinion Research* 22 (4): 535-551.

107. Yan, Ting, Curtin, Richard in Jans, Matthew. 2010. Trends in Income Nonresponse Over Two Decades. *Journal of Official Statistics* 26 (1): 145-164.
108. Widaman, Kieth F. 2006. Missing data: What to do with or without them. *Monographs of the Society for Research in child development* 71 (3): 42-64.
109. *Zakon o varstvu osebnih podatkov* (ZVOP-1-UPB1). Ur. l. RS 94/2007 (16. oktober 2007).

7 PRILOGA

7.1 Priloga 1

Tabela 7.1: Kategorije izidov ankete.

Kategorija	**EHIS 2007**
I= V celoti izvedene ankete (ang. Complete Interviews)	2117
P=Delno izvedene ankete (ang. Partial Interviews)	1
R=Nesodelovanje (ang. Refusal and break off)	799
NC=Brez stika (ang. Non Contact)	202
O=Drugo (ang. Other)	0
UH=Neznano gospodinjstvo (ang. Unknown Household)	0
UO=Drugo neznano (ang. Unknown Other)	0

Vir: EHIS 2007.

Tabela 7.2: Izračun stopenj odgovora in stopenj sodelovanja v EHIS 2007.

STOPNJE ODGOVORA	
Stopnja odgovora 1	
I/(I+P) + (R+NC+O) + (UH+UO)	0,679
Stopnja odgovora 2	
(I+P)/(I+P) + (R+NC+O) + (UH+UO)	0,679
Stopnja sodelovanja 1	
I/(I+P)+R+O)	0,726
Stopnja sodelovanja 2	
(I+P)/((I+P)+R+0))	0,726
Stopnja sodelovanja 3	
I/((I+P)+R))	0,726

Stopnja sodelovanja 4	
(I+P)/((I+P)+R))	0,726

Vir: AAPOR 2011

Tabela 7.3: Izračun stopenj zavračanja in stopenj stika v EHIS 2007.

Kategorija	**EHIS 2007**
Stopnje zavračanja	
Stopnja zavračanja 1	
R/((I+P)+(R+NC+O) + UH + UO))	0,256
Stopnja zavračanja 2	
R/((I+P)+(R+NC+O)) + e(UH + UO))	0,256
Stopnja zavračanja 3	
R/((I+P)+(R+NC+O))	0,256
Stopnje stika	
Stopnja stika 1	
(I+P)+R+O / (I+P)+R+O+NC + (UH + UO)	0,935
Stopnja stika 2	
(I+P)+R+O / (I+P)+R+O+NC + e (UH + UO)	0,935
Stopnja stika 3	
(I+P)+R+O / (I+P)+R+O+NC	0,935

Vir: AAPOR 2011.

7.2 Priloga 2

Za izračun uteži smo uporabili naslednje populacijske totale oz. populacijske margine:

a) **Tabela 7.4:** Tip naselja.

	N
nekmečka naselja z manj kot 2000 prebivalci	514813
kmečka naselja z manj kot 2000 prebivalci	407145
naselja z 2000 do 10000 prebivalci	281402
naselja z 10000 do 100000 prebivalci	233271
Maribor	82300
Ljubljana	225854
skupaj	**1744785**

Vir: Centralni register prebivalstva 2007.

b) **Tabela 7.5:** Spol x petletni starostni razredi.

Spol	Star. razredi	N
moški	15-19	59199
	20-24	70889
	25-29	81080
	30-34	80032
	35-39	76052
	40-44	80379
	45-49	79496
	50-54	81052
	55-59	71465
	60-64	49555
	65-69	45485
	70-74	36138
	75-79	26100
	80-84	13056
	85+	6136
ženske	15-19	55369
	20-24	64749
	25-29	73113
	30-34	73098
	35-39	70360
	40-44	77455
	45-49	75326
	50-54	76819
	55-59	69373
	60-64	53080
	65-69	53076
	70-74	49245
	75-79	44637
	80-84	32512
	85+	20459
skupaj		**1744785**

Vir: Centralni register prebivalstva 2007.

c) **Tabela 7.6:** Statistična regija.

Statistična regija	N
Pomurska	105826
Podravska	280190
Koroška	63385
Savinjska	224172
Zasavska	39465
Spodnjeposavska	60647
Jugovzhodna Slovenija	120305
Osrednjeslovenska	435052
Gorenjska	171322
Notranjsko-kraška	44989
Goriška	104340
Obalno-kraška	95092
skupaj	**1744785**

Vir: Centralni register prebivalstva 2007.

d) **Tabela 7.7:** Dosežena stopnja izobrazbe.

	N
Brez šolske izobrazbe oz. nepopolna osnovna izobrazba, 1 – 3	17448
Nepopolna osnovna izobrazba, 4 – 7 razredov.	59323
Osnovna izobrazba.	394321
Nižja ali srednja poklicna izobrazba.	437941
Srednja strokovna izobrazba.	434451
Srednja splošna izobrazba.	104687
Višja strokovna izobrazba, višješolska izobrazba.	90729
Visoka strokovna izobrazba.	52344
Visoka univerzitetna izobrazba.	130859
Specialistična povisokošolska izobrazba, magisterij, doktorat	22682
skupaj	**1744785**

Vir: Anketa o delovni sili 2007.

7.3 Priloga 3

Tabela 7.8: Test multikolinearnosti - koeficienti.

Coefficients[a]

Model		Collinearity Statistics	
		Tolerance	VIF
1	spol Spol anketiranca	*.952*	*1.051*
	star_sk Starost rekodirana v 3 razrede	*.551*	*1.814*
	st_urb Stopnja urbanizacije	*.968*	*1.033*
	zak_stan Formalni zakonski stan.	*.629*	*1.591*
	zap_status Trenutni zaposlitveni status	*.691*	*1.446*
	dos_izob Dosezena stopnja izobrazbe	*.723*	*1.383*
	doh_gosp Skupen dohodek gospodinjstva - rekodiran	*.715*	*1.398*

a. Dependent Variable: nerespondent nerespondent ne glede na nacin anketiranja

Tabela 7.9: Test multikolinearnosti.

Collinearity Diagnostics[a]

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions							
				(Constant)	spol Spol anketiranca	star_sk Starost rekodirana v 3 razrede	st_urb Stopnja urbanizacije	zak_stan Formalni zakonski stan.	zap_status Trenutni zaposlitveni status	dos_izob Dosezena stopnja izobrazbe	doh_gosp Skupen dohodek gospodinjstva - rekodiran
1	1	7.162	1.000	.00	.00	.00	.00	.00	.00	.00	.00
	2	.371	4.394	.00	.00	.00	.00	.01	.24	.03	.11
	3	.135	7.288	.00	.01	.02	.03	.13	.51	.00	.28
	4	.120	7.725	.00	.03	.03	.46	.09	.00	.07	.01
	5	.094	8.734	.00	.76	.03	.16	.01	.00	.00	.01
	6	.066	10.378	.00	.00	.00	.00	.23	.03	.63	.46
	7	.037	13.991	.01	.01	.76	.03	.53	.17	.16	.01
	8	.015	21.503	.99	.18	.16	.32	.00	.04	.11	.13

a. Dependent Variable: nerespondent nerespondent ne glede na nacin anketiranja