

UNIVERZA V LJUBLJANI
FAKULTETA ZA DRUŽBENE VEDE

Nino Zajc

Kombiniranje različnih načinov anketiranja v uradnih anketah

Diplomsko delo

Ljubljana, 2011

UNIVERZA V LJUBLJANI
FAKULTETA ZA DRUŽBENE VEDE

Nino Zajc

Mentor: red. prof. dr. Vasja Vehovar

Somentor: asist. Jernej Berzelak

Kombiniranje različnih načinov anketiranja v uradnih anketah

Diplomsko delo

Ljubljana, 2011

Zahvala

Hvala mentorju red. prof. dr. Vasji Vehovar in somentorju asist. Jerneju Berzelaku za strokovne nasvete in podporo pri pisanju diplomske naloge.

Hvala mami, očetu in sestri za podporo v času študija.

Kombiniranje različnih načinov anketiranja v uradnih anketah

Izbira raziskovalnega načrta, ki zagotavlja optimalnost med kakovostjo podatkov (napakami) in stroški, predstavlja velik izziv sodobnega anketnega raziskovanja. Optimalen izbor posebej otežuje upadanje sodelovanja. Vedno več raziskovalcev se zato odloča za uporabo kombiniranih načinov anketiranja, s katerimi lažje optimiziramo napake in stroške. V diplomski nalogi želimo odgovoriti na raziskovalno vprašanje, kako se v raziskavi o uporabi informacijsko komunikacijske tehnologije v gospodinjstvih 2008 (IKT_GOS), ki jo letno izvaja Statistični urad Republike Slovenije, spreminja odnos med srednjo kvadratno napako in stroški. Pri tem upoštevamo, da se stroški in napake med spremenljivkami močno razlikujejo. Pri večini spremenljivk se pokaže, da je optimalna izbira telefonski način, pri ostalih pa z večanjem vzorca to postane terenski način. S simulacijo nagrajevanja se izkaže, da je pri nekaterih spremenljivkah optimalno vključiti tudi nagrade.

Ključne besede: kombiniranje načinov, srednja kvadratna napaka, optimizacija, anketni stroški, nagrade.

Mixed mode data collection in official surveys

The selection of a survey design, which ensures an optimal quality of data (errors) and costs, represents a major challenge of the modern survey research. Such selection is difficult to make, due to the decline in participation. Hence a majority of researchers choose to use combined methods of the survey to optimize errors and costs better. In our thesis we want to answer the research question about how - in the survey on the use of information communication technology in households 2008 (IKT_GOS), which is annually performed by the Statistical Office of the Republic of Slovenia - the relationship between the mean squared error (MSE) and cost changes. Costs and errors vary between the variables. At most variables the most optimal choice is the telephone mode, while at the others, by increasing the sample, the most optimal way becomes the personal mode. With the simulation of rewarding an individual way the results show that, at some variables, the most optimal choice is the mode with incentives.

Keywords: mix mode, mean squared error, optimization, survey costs, incentives.

Kazalo

1	UVOD.....	8
2	NAPAKE V ANKETAH.....	10
2.1	VZORČNE NAPAKE.....	11
2.2	NEVZORČNE NAPAKE.....	11
2.2.1	NAPAKA POKRITJA	11
2.2.2	MERSKA NAPAKA.....	12
2.2.3	NAPAKA NEODOGVORA	13
2.2.4	NAPAKA OPREDELITVE	14
2.2.5	NAPAKA PRI PROCESIRANJU PODATKOV	14
3	METODE ZBIRANJA PODATKOV	15
3.1	KLASIFIKACIJA METOD ZBIRANJA PODATKOV	16
3.1.1	TERENSKO ANKETIRANJE.....	17
3.1.2	TELEFONSKO ZBIRANJE PODATKOV	19
3.1.3	POŠTNO ANKETIRANJE	21
3.1.4	SPLETNO ANKETIRANJE	22
3.2	IZBIRA USTREZNE METODE	24
3.2.1	ODVISNOST OD RAZISKOVANE POPULACIJE	24
3.2.2	ODVISNOST OD VIROV	25
3.2.3	ODVISNOST OD VPRAŠALNIKA	25
3.3	VIDIK STROŠKOV	25
4	KOMBINIRANJE NAČINOV ANKETIRANJA.....	27
4.1	RAZLOGI ZA UPORABO KOMBINIRANIH NAČINOV	28
4.2	KLASIFIKACIJA KOMBINIRANIH NAČINOV	29
4.2.1	KOMBINIRANJE NAČINOV V FAZI ZAČETNEGA KONTAKTIRANJA	30
4.2.2	KOMBINIRANJE NAČINOV V FAZI ZBIRANJA PODATKOV	31
4.2.3	KOMBINIRANJE NAČINOV V FAZI OPOMNIKOV	34
4.2.4	DRUGE OBLIKE KOMBINIRANJA NAČINOV	35
4.3	DILEME KOMBINIRANJA NAČINOV	35
4.4	KLASIFIKACIJA ANKET SURS.....	35
5	EMPIRIČNI DEL	43
5.1	OPIS RAZISKAVE.....	43

5.2	OPIS VZORCA.....	43
5.3	PRILAGODITEV VZORCA.....	44
5.4	IZBIRA SPREMENLJIVK.....	45
5.5	METODA, ANALIZA IN INTERPRETACIJA	46
5.6	OMEJITVE.....	56
6	ZAKLJUČEK.....	57
7	LITERATURA	60
	PRILOGI	64
	PRILOGA A: TABELE OCENITVE STROŠKOV POSAMEZNEGA NAČINA...	64
	PRILOGA B: SIMULACIJE SPREMENLJIVK	66

Kazalo slik

Slika 2.1:	Skupna anketna napaka.....	10
Slika 3.1:	Simulacija stroškov in velikosti vzorca.	26
Slika 3.2:	Simulacija stroškov glede na način in velikost vzorca ob predpostavki o 100-odstotni stopnji sodelovanja.	27
Slika 4.1:	Ločevanje faze kontaktiranja in zbiranja podatkov.	29
Slika 5.1:	Simulacija spremenljivke "pravica do izmenjave zakonsko zaščitene vsebine" brez nagrade in z nagrado	52
Slika 5.2:	Simulacija spremenljivke "uporaba mobilnega telefona" brez nagrade in z nagrado	53
Slika 5.3:	Simulacija spremenljivke "dostop do interneta" brez nagrade in z nagrado.....	53
Slika 5.4:	Simulacija spremenljivke "pogostost uporabe interneta" brez nagrade in z nagrado	54

Kazalo tabel

Tabela 3.1:	Klasifikacija metod anketnega zbiranja podatkov glede na vrsto kontakta, vključenost računalniške tehnologije in vpletenosti anketarja	16
Tabela 3.2:	Razmerja stroškov med CATI- in CAPI-načinom anketiranja	27
Tabela 4.1:	Klasifikacija načinov anketiranja po de Leeuw (2005)	38

Tabela 4.2: Klasifikacija raziskav SURS po dopolnitvi klasifikacije Edith de Leeuw (2005)	39
Tabela 4. 3: Podrobni podatki o raziskavah SURS	40
Tabela 5.1: Izbira in rekodacija spremenljivk	45
Tabela 5.2: MSE in stroški, telefonsko anketiranje.	49
Tabela 5.3: MSE in stroški za številske spremenljivke, telefonsko anketiranje	49
Tabela 5.4: MSE in stroški, terensko anketiranje	50
Tabela 5.5: MSE in stroški za številske spremenljivke, terensko anketiranje	50

1 Uvod

Anketno raziskovanje se kot metoda zbiranja podatkov uporablja na številnih in raznolikih področjih. Zaradi vse večje ponudbe metod zbiranja podatkov se raziskovalci soočajo s problemom, katero metodo izbrati, da bi pridobili čim kakovostnejše podatke. Hkrati s tem se vse več organizacij in uradov sooča s finančnimi omejitvami, ki vplivajo tudi na kvaliteto zbranih podatkov.

Biemer in Lyberg (2003) navajata, da v sodobnih družbah zaradi poplave anket podjetij, zavarovalnic in ostalih organizacij nastaja vtis, da je sodelovanje v anketah vse prej kot koristno, nemalokrat pa odvečno in celo mučno. Kolateralna škoda nastaja v tem primeru tudi za organizacije, ki se strokovno ukvarjajo z anketnim raziskovanjem, vključno z nacionalnimi statističnimi uradi. Tudi v Sloveniji se Statistični urad Republike Slovenije (SURS) sooča z naraščajočim problemom nesodelovanja oseb, vse pogosteje pa tudi z grobimi zavrnitvami.

Eden od pristopov za reševanje problema naraščajočih stroškov in večanja stopenj nesodelovanja je kombiniranje načinov anketiranja. De Leeuw (2005) navaja, da se zato danes vse več raziskovalcev odloča za kombiniranje načinov anketiranja (*angl. mixed mode*). S tem pristopom se združi prednosti različnih metod in hkrati kompenzira njihove slabosti. Glavni razlog kombiniranja načinov je, kot že omenjeno, optimizacija stroškov in napak. Za kombiniranje načinov se lahko odločimo v vseh fazah raziskovalnega procesa, zato obstaja veliko kombinacij. Seveda pa lahko v primeru, ko se raziskovalci nepremišljeno odločijo za kombiniranje, nastanejo tudi negativni učinki na kakovost podatkov in na stroške raziskave (de Leeuw 2005).

V diplomskem delu se najprej (poglavje 2 in 3) osredotočamo na pregled osnovnih konceptov napak v anketah (*angl. survey errors*) in klasifikacije metod zbiranja podatkov ter izvedemo simulacijo stroškov telefonskega in terenskega zbiranja podatkov. V četrtem delu diplomske naloge bomo predstavili kombinirane načine anketiranja, razširjeno klasifikacijo po de Leeuw (2005), dileme kombiniranja ter kombiniranje načinov v raziskavah SURS.

V osrednjem (empiričnem) delu bomo na podatkih raziskave IKT_GOS 2008 odgovorili na raziskovalno vprašanje:

Kako se spreminja odnos med stroški in napakami med načini anketiranja posamezne spremenljivke z večanjem vzorca?

V ta namen bomo izbrali 13 spremenljivk različnih merskih in vsebinskih tipov in primerjali optimalno izbiro metode za posamezno spremenljivko. Predpostavljamo, da se ocena za pristranskost z večanjem vzorca ne spreminja, spreminja se le vzorčna varianca, ki se z večanjem vzorca manjša. Zaradi naraščajočega trenda daril in nagrad (*angl. incentives*) za anketirance bomo naredili tudi simulacijo z ocenami stopenj odgovorov, ki bi jih dobili, če bi vsakemu anketirancu (Lozar Manfreda in drugi 2009, 7), ki odgovori na telefonsko anketo, ponudili 10 €, tistemu, ki odgovori na terensko anketo, pa 5 € nagrade. V raziskovalnem delu bomo torej primerjali štiri načine.

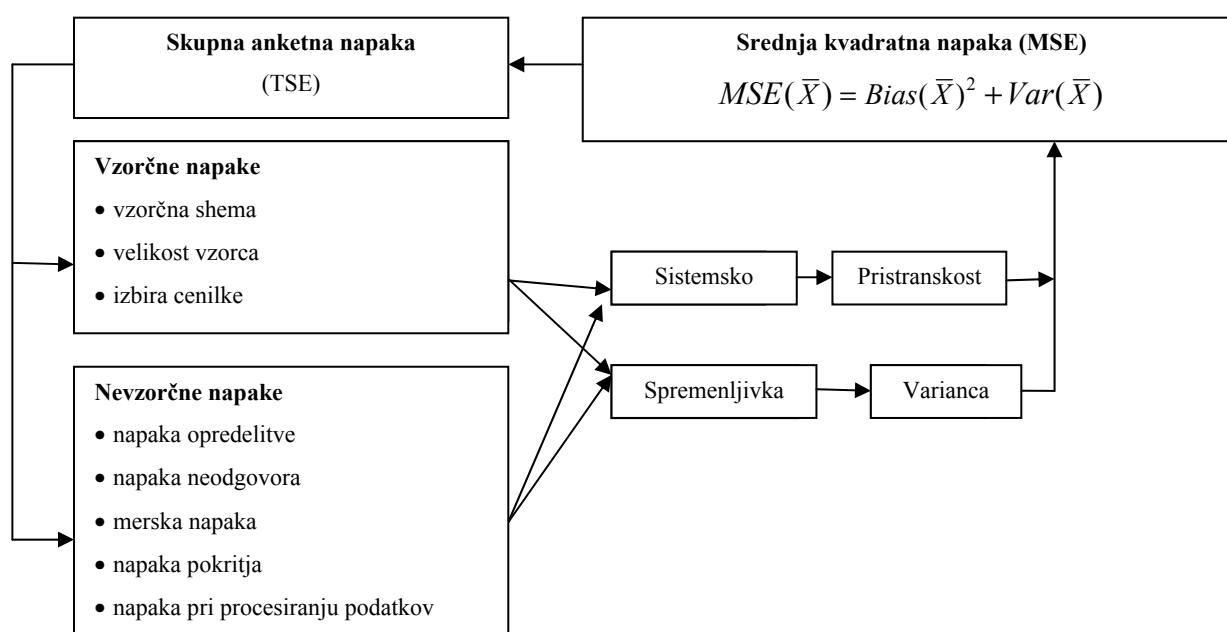
V zadnjem delu (poglavje 6) bomo ugotovitve diplomske naloge povzeli ter podali predloge za raziskovanje obravnavane problematike.

2 Napake v anketah

V literaturi lahko najdemo različne klasifikacije anketnih napak (Biemer in Lyberg 2003; Groves 1989, 2005; Groves in ostali 2004; Lessler in Kalsbeek 1992; Weisberg 2005). V diplomski nalogi bomo obravnavali klasifikacijo vzorčnih (*angl. sampling*) in nevzorčnih (*angl. non-sampling*) anketnih napak na osnovi Biemer in Lyberg (2003). Ne glede na klasifikacijo anketnih napak je vsem klasifikacijam skupno, da je vsota vseh anketnih napak označena kot skupna anketna napaka (*angl. total survey error – TSE*), ki predstavlja pričakovane razlike med ocenami na vzorcu in pravimi populacijskimi vrednostmi (Groves, 1989). Te razlike so lahko slučajne ali pa sistematčne. Biemer in Lyberg (2003) skupno anketno napako definirata kot vsoto vzorčnih in nevzorčnih anketnih napak (Slika 2.1). V diplomskem delu bomo operirali s srednjo kvadratno napako (*angl. mean square error – MSE*) kot mero skupne anketne napake, ki jo izrazimo v obliki (Deming 1950):

$$MSE(\bar{X}) = Bias(\bar{X})^2 + Var(\bar{X})$$

Slika 2.1: Skupna anketna napaka



Vir: Prirejena slika po Biemer (2010, 822).

2.1 Vzorčne napake

Vzorčne napake izhajajo iz dejstva, da v seznamu ni vseh enot ciljne populacije, ampak samo vzorec le-teh. Zato informacije, zbrane od teh enot v vzorcu, morda ne odražajo popolnoma informacij, ki bi jih dobili, če bi jih zbrali iz cele populacije. Ta razlika je vzorčna napaka. Vzorčna teorija predpisuje formule za ocenjevanje vzorčnih napak, ki se jih uporablja pri verjetnostnih vzorcih (verjetnostni vzorci so tisti, pri katerih imajo vse enote v ciljni populaciji znano in neničelno verjetnost, da so izbrane v vzorec; da je vzorec verjetnosten, je tudi običajna predpostavka pri izračunavanju vzorčnih napak). Če pa vzorec ni verjetnosten, ne moremo uporabiti običajne teorije, poročamo lahko le o približni oceni, ki jo dobimo iz obravnave le dela enot (Statistični urad RS 2003). Vzorčno napako se zmanjšuje s povečanjem velikosti vzorca, ker se s tem zmanjšuje verjetnost, da bo vzorec statistično značilno različen od populacije (Lozar Manfreda in drugi 2000, 1038). V praksi je velikost vzorca največkrat omejena s stroški.

2.2 Nevzorčne napake

Nevzorčne napake se lahko pojavijo v vseh fazah raziskovalnega procesa. Raziskovalci poskušajo nevzorčno napako čim bolj kontrolirati, saj njena stopnja vpliva na točnost podatkov (Biemer in Lyberg 2003, 34). Raziskovalci delijo nevzorčne anketne napake na pet osnovnih tipov napak.

2.2.1 Napaka pokritja

Napake pokritja (*angl. coverage errors*) nastanejo zaradi neskladja med ciljno populacijo in okvirom, kjer je ciljna populacija tista skupina, ki nas zanima za raziskovanje in jo obravnavamo pri nadaljnjih statističnih interpretacijah. Okvir predstavljajo enote, od katerih zberemo informacije, ki nas zanimajo (Biemer in Lyberg 2003, 40). Raziskovalci največkrat zbirajo podatke le od nekaterih enot (vzorec), pri popisih prebivalstva pa se uporablja polno zajetje enot (anketiramo vse enote populacije). Napake pokritja skušajo raziskovalci preseči s pomočjo uporabe ustreznih metod uteževanja (Lozar Manfreda in drugi 2000).

Neujemanje med okvirom in ciljno populacijo je vir napak pokritja. Napake pokritja vključujejo nadpokritje, podpokritje in napačno razvrščanje statističnih enot (Statistični urad RS 2003, 7).

- Nadpokritje (*angl. overcoverage*) je v modernem raziskovanju še nedorečen pojem. Nekateri so mnenja, da gre pri nadpokritju za problem neustreznih enot, drugi pa so mnenja, da gre za problem podvojenih enot, zato bomo pojem nadpokritja obravnavali kot skupek neustreznih in podvojenih enot. Nadpokritje se pojavi, ko se več kot ena enota v vzorčnem okviru nanaša na isto enoto v ciljni populaciji (Biemer in Lyberg 2003, 65). Primer nadpokritja so podvojene enote (*angl. duplicate units*), npr. če ima pri anketiranju po mobilnih telefonih ista oseba več različnih telefonskih števil (SIM-kartice) (Belak 2008, 61). K nadpokritju prištevamo tudi neustrezne enote (*angl. uneligible units*), ki se jih večinoma odkrije med procesom zbiranja podatkov (npr. tujci, osebe brez volilne pravice itd.). Neustrezne enote (*angl. uneligible units*) se nanašajo tudi na tuje enote (*angl. foreign units*), ki se sicer nahajajo v okviru, vendar ne sodijo v ciljno populacijo, npr. brezposelni v raziskavi o podrobnostih dela na delovnem mestu ali mlajši od 18 let v volilni raziskavi (Kalton in Vehovar 2001, 72). Število neustreznih enot je potrebno vnaprej predvideti in oceniti njihov delež ter potem ustrezno povečati začetni vzorec.
- Podpokritje (*angl. undercoverage*) se nanaša na enote, ki niso vključene v okvir, vendar bi morale biti (npr. hočemo anketirati vsa gradbena podjetja, a zaradi neažurnosti podatkov nekatera spustimo).
- Napačno razvrščanje enot (*angl. misclassification*) se nanaša na razvrščanje statističnih enot v napačne razrede glede na dano klasifikacijo (Statistični urad RS 2003, 8), npr. podjetje je razvrščeno pod trgovine namesto v industrijo.

2.2.2 Merska napaka

Merske napake (*angl. measurement errors*) so napake, ki se pojavijo v fazi zbiranja podatkov. Za številne raziskave je merska napaka ena izmed največjih virov anketnih napak (Biemer in Lyberg 2003, 41). Biemer in Lyberg (2003, 42) opredelita napako kot razliko med dejansko in

zabeleženo vrednostjo. Groves (1989) pravi, da je merska napaka rezultanta anketnega vprašalnika oziroma načina, kako je posamezno vprašanje respondentu postavljeno.

V splošnem lahko vire merskih napak razdelimo (Biemer in Lyberg 2003, 36) na:

- anketarja (vpliv anketarja na odgovore anketirancev)
- anketne instrumente (slabo sestavljeno vprašanje ali celotni vprašalnik)
- anketiranja (nesodelovanje anketiranja, nezanimanje, psihološki dejavniki)
- način zbiranja podatkov (vsak način ima svoje prednosti in pomanjkljivosti v razmerju do merske napake)

2.2.3 Napaka neodgovora

Napake zaradi neodgovora (*angl. nonresponse errors*) dobimo, če podatkov od enot, ki so za raziskavo ustrezni in so bili vključeni v vzorec, ne dobimo. Največja težava pri neodgovoru je nevarnost, da se elementi, ki so v anketni raziskavi sodelovali, razlikujejo od ostalih ustreznih elementov, ki v raziskavi niso sodelovali (Kalton in Vehovar 2001, 78).

Ločimo dve skupini neodgovorov.

- Neodgovor enote (*angl. unit nonresponse*) je takrat, kadar od ustrezne enote ne dobimo nobene informacije. Neodgovor enote delimo na kategorije: zavrnitev sodelovanja (*angl. refusal*), nezmožnost sodelovanja, odsotnost anketiranja, nezmožnost lociranja ustreznega anketiranja, nedostopnost anketiranja, izgubljeni ali poškodovani vprašalniki (Kalton in Vehovar 2001, 80).
- Neodgovor spremenljivke (*angl. item nonresponse*) je takrat, kadar ustrezna enota sicer odgovarja na vprašanja, vendar na nekatera vprašanja noče odgovoriti. Vzroki za neodgovor spremenljivke so lahko različni: anketiranci morda ne poznajo odgovora na vprašanje, anketiranci posredno zavrnejo odgovor na vprašanje, anketar lahko vprašanje pozabi ali preskoči, odgovor je sicer lahko vpisan v vprašalnik, vendar je izločen, ker ni bil konsistenten z drugimi odgovori, nekateri odgovori se v procesu obdelave podatkov izgubijo (Kalton in Vehovar 2001, 88).

2.2.4 Napaka opredelitve

Do napake opredelitve pride v fazi merjenja (Biemer 2010, 822). Pri tej napaki gre za uporabo slabih merskih instrumentov (npr. z vprašanjem merimo drugo stvar namesto tiste, ki bi jo radi merili). Prisotnost te napake povzroči problem veljavnosti. Pri veljavnosti je pomembna predvsem vsebinska veljavnost, kjer je potrebno z mersko lestvico pokriti vsebino merjenega koncepta oziroma merjene spremenljivke. Druga izmed pomembnih veljavnosti je kriterijska veljavnost, pri kateri gre za uporabo preizkušenih merskih instrumentov. Pri veljavnosti konstrukta pa gre za povezavo med dvema teoretskima spremenljivkama oziroma za korelacijo med merskimi instrumenti teh dveh spremenljivk.

2.2.5 Napaka pri procesiranju podatkov

Napake pri procesiranju podatkov nastanejo izključno s strani raziskovalca oziroma osebe, ki obdeluje podatke. Napake največkrat nastanejo pri šifriranju, vnosu, urejanju, uteževanju in tabeliranju podatkov.

3 Metode zbiranja podatkov

Sistematično in množično anketno raziskovanje se je razvilo v tridesetih letih prejšnjega stoletja. Sprva so se preko standardiziranih anket izvajali državni popisi prebivalstva. Tem so sledila družboslovna raziskovanja, s katerimi še danes pojasnjujemo družbene pojave. Z vzponom kapitalizma so se ankete začele uporabljati tudi v tržnem raziskovanju. Procesi anketnega raziskovanja se spreminjajo skladno s tehnološkim razvojem. V tridesetih letih prejšnjega stoletja so raziskovalci poleg že znanega terenskega načina zbiranja podatkov začeli uporabljati tudi telefon. Ker je bilo uporabnikov telefona takrat še malo, so se pogosto soočali z napako nepokritja (Nathan 2001, 8). V poznih šestdesetih letih se je zaradi velike pokritosti v zahodni Evropi in ZDA (88 % gospodinjstev je imelo telefon) telefonsko raziskovanje začelo uporabljati kot primaren način zbiranja podatkov (Massey in Botman v Nathan 2001, 8).

Z razvojem računalniške tehnologije v 80. letih so se v procesu anketnega zbiranja podatkov zvrstile številne spremembe na že obstoječih načinih zbiranja in obdelave podatkov. Poštne ankete danes razbiramo in kodiramo z ICR-tehnologijo (*angl. Intelligent Character Recognition*) ali pa pošljemo vprašalnik na disketi, CD-ju ali drugem prenosnem mediju. Terenski anketarji so začeli uporabljati računalnike za anketiranje, v katere vpisujejo odgovore anketiranca. Računalniško zbiranje podatkov na terenu pa je naredilo še korak naprej z računalniškim samoanketiranjem, kjer ima anketiranec možnost poslušati ali prebrati vprašanje ter nato tudi odgovoriti preko računalnika.

V telefonskem načinu zbiranja podatkov se je uvedlo računalniško podprto telefonsko anketiranje ter z njim interaktivni glasovni odgovor (*angl. Interactive Voice Response – IVR*). Programska oprema pri računalniško podprtem telefonskem anketiranju omogoča sprejetje in preverjanje odgovorov, ki jih vnese anketar. Z vzponom interneta v 90. letih se je uvedlo računalniško podprto spletno anketiranje. Razvoj metod zbiranja podatkov je bil torej hiter, a vseeno je vsem metodam skupno, da se je tehnologija razvijala v okviru klasičnih načinov zbiranja podatkov, osebnega, telefonskega in poštne anketiranja. Zato je na tehnologijo iz sociološko raziskovalnega vidika potrebno gledati in jo razumeti skozi sedem ključnih pogledov: tehnologija kot zamenjava, podpora, omogočanje, uspešnost, nadzor, informiranje in kot nadomestek tehnologije (Couper 2002, 17).

Skupaj z vplivom tehnologije na načine zbiranja podatkov je potrebno upoštevati tudi kakovost zbranih podatkov. Biemer in Lyberg (2003) v knjigi *Introduction to Survey Quality* obravnavata različne načine zbiranja podatkov in z njimi povezane anketne napake, ki jih bomo opisali in klasificirali v drugem delu diplomske naloge.

3.1 Klasifikacija metod zbiranja podatkov

Tabela 3.1: Klasifikacija metod anketnega zbiranja podatkov glede na vrsto kontakta, vključenost računalniške tehnologije in vpletenosti anketarja

	Visoka vključenost anketarja		Nizka vključenost anketarja	
	Ankete na papirju	Računalniške ankete	Ankete na papirju	Računalniške ankete
Neposreden kontakt	Terenska anketa z vprašalnikom na papirju (PAPI)	CAPI	Dnevnik	CASI, ACASI, elektronski dnevniki
Posreden kontakt	Telefonska anketa z vprašalnikom na papirju (PAPI)	CATI	Poštna anketa, anketa po faksu	TDE, WEB, DBM, EMS, VRE

Razlaga kratic: **ACASI** – glasovno računalniško podprto samoanketiranje (*angl. audio computer- assisted self- interviewing*), **CASI** – računalniško podprto samoanketiranje (*angl. computer-assisted self-interviewing*), **CADE** – (*angl. computer-assisted data entry*), **CAPI** – računalniško podprto osebno anketiranje (*angl. computer assisted personal interviewing*), **CASI** – računalniško podprto samoanketiranje (*angl. computer assisted self-interviewing*), **CATI** – računalniško podprto telefonsko anketiranje (*angl. computer assisted telephone interviewing*), **DBM** – anketa na prenosnem mediju poslana po pošti (*angl. disk by meil*), **EMS** – anketa po elektronski pošti (*angl. electronic meil survey*), **PAPI** – anketa na papirju (*angl. pen-and-paper interviewing*), **T- ACASI** – ACASI preko telefona, **VRE** – prepoznavanje glasu (*angl. Voice Recognition Entry*)

Vir: Prirejena tabela po Biemer in Lyberg (2003, 189).

V Tabeli 3.1 so predstavljeni različni načini zbiranja podatkov glede na vrsto kontakta, vpletenost računalniške tehnologije in vpletenosti anketarja. Biemer in Lyberg (2003, 189) ločita med dvema vrstama anket glede na medij: anketo na papirju in anketo na računalniku. Ločita tudi med anketami z visoko in nizko vključenostjo anketarja, kjer visoka vključenost

pomeni prisotnost napak, ki nastanejo zaradi anketarjevega učinka, nizka vključenost pa pomeni, da je učinek anketarja zelo majhen. Biemer in Lyberg (2003, 189) ločita tudi glede na vrsto kontakta, ki je lahko neposreden ali posreden. Neposreden kontakt pomeni prisotnost anketarja, medtem ko posreden pomeni, da anketar ni prisoten oziroma je prisoten, vendar ne fizično (npr. preko telefona). V sodobnem raziskovanju se raziskovalci najpogosteje odločajo za uporabo računalniških anket, kjer je vprašalnik posredovan preko računalnika (CAPI, CATI, CASI, ACASI in WEB). Računalniško posredovane ankete imajo nekatere prednosti in slabosti v primerjavi s klasičnimi načini. Glavne prednosti so, da je zbiranje podatkov hitrejše, cenejše, zabeleženi podatki gredo neposredno v bazo podatkov, z računalniško posredovanim vprašalnikom zmanjšamo učinek anketarja in omogočimo izvedbo kompleksnih vprašalnikov. Couper (v Biemer in Lyberg 2003, 195) trdi, da prisotnost standardiziranih vprašalnikov v računalniško posredovanih anketah omogoča večjo obvladljivost merskih napak. V nadaljevanju diplomske naloge bomo obravnavali pet najpogosteje uporabljenih načinov zbiranja podatkov: terensko, telefonsko, pošno, spletno in kombinirano anketiranje.

3.1.1 Terensko anketiranje

Terensko anketiranje je najstarejši način, ki se ne nanaša na tehnološke inovacije in zagotavlja najvišjo stopnjo komunikacije in interakcije med anketarjem in anketirancem. Zaradi dobre kakovosti podatkov in visoke stopnje odgovora je ta način še vedno eden izmed najpogosteje uporabljenih načinov zbiranja podatkov na različnih področjih raziskovanja. Raziskovalci se pri raziskovanju občutljivih tem (spolnost, alkoholizem, deviantnost itd.) ne odločajo za ta način zbiranja podatkov, ker tvegajo visoko vrednost merske napake.

Z razvojem IKT v devetdesetih letih se je razvil tudi sistem računalniškega anketiranja (*computer assisted interviewing – CAI*). Najprej so CAI uporabljali v telefonskem anketiranju, kasneje so sistem CAI začeli uporabljati tudi raziskovalci pri terenskem osebnem anketiranju (*angl. computer assisted personal interviewing – CAPI*). Danes je ta način eden izmed najpogostejših načinov zbiranja podatkov. Anketarji v CAPI-načinu nimajo več papirnatih anket, temveč imajo anketo na prenosnem računalniku, kamor tudi beležijo odgovore anketiranca, ti odgovori pa gredo neposredno v zbirno bazo.

Groves in Tortora (1998) trdita, da nam uporaba računalniškega zbiranja in vodenja podatkov v osnovi omogoča, da se izognemo nekaterim napakam.

Terensko anketiranje v splošnem velja za najdražji način zbiranja podatkov. Ker se podatke zbira na terenu, morajo anketarji obiskati respondenta na domu, v službi ali na javnem mestu. Potni stroški pa danes v raziskovanju predstavljajo velik delež končnih stroškov. Poleg visokih potnih stroškov in stroškov usposabljanja anketarjev se pri osebni anketiranju srečujemo tudi s problemom vpliva anketarja.

Anketarji pri osebni načinu zbiranja podatkov nehoti vplivajo na odgovore pri občutljivih vprašanjih. Le-ti lahko postanejo proizvajalci fenomena družbeno zaželenih odgovorov in pristranskosti ali celo neodgovora spremenljivke. Na odgovor respondenta tako lahko vpliva tudi anketarjevo obnašanje, hitrost branja vprašanja, delovne procedure anketarja, beleženje odgovora respondenta ter način postavljanja vprašanja in njegovega zaključevanja. Da bi se zmanjšal vpliv anketarja, je potrebno standardizirati njegovo obnašanje med anketiranjem in tudi obnašanje v fazi začetnega kontakta z izobraževanjem (Biemer in Lyberg 2003, 191).

Kljub anketarjevemu vplivu in socialno zaželenim odgovorom je terensko anketiranje še vedno eden izmed najbolj fleksibilnih načinov, saj omogoča izvedbo kompleksnih in dolgih anket, ki lahko trajajo tudi več kot uro. Anketar lahko z uporabo različnih tehnik prepriča respondenta k sodelovanju, lahko opazuje respondenta med anketiranjem ter zagotavlja, da ni anketa izpolnjena s prisotnostjo ali pomočjo drugih ljudi. Čeprav lahko anketar nadzira prisotnost drugih oseb, ima še vedno veliko vlogo respondent. Anketarji lahko pri osebni načinu pomagajo respondentu pri razumevanju vprašanja bolj kot pri drugih načinih, ker lahko za bolj občutljiva vprašanja ponudijo kartice, na katere respondent odgovori (*angl. response card*). Ena izmed glavnih prednosti osebnega anketiranja je visoka pokritost splošne populacije, saj država vodi centralni register prebivalstva, ki je dostopen Statističnemu uradu Republike Slovenije.

Velik problem predstavlja tudi ponarejanje anket s strani anketarjev. Tako ravnanje anketarjev je poznano vsem organizacijam, ki se ukvarjajo z zbiranjem podatkov, in ga je zelo težko zaznati. Anketar lahko ponaredi celotno anketo ali pa spusti določena vprašanja z namenom, da si prihrani čas. Na ta fenomen ponarejanja podatkov sta opozorila Biemer in Stokes (v

Biemer in Lyberg 2003, 191). Mogoče ga je odkriti s ponovnim anketiranjem, preverjanjem anketiranja ali z analizo skupnega odziva na anketo. Največji problem osebnega anketiranja predstavljajo torej stroški, ki so povezani s samo izvedbo anketiranja. Ker morajo anketarji potovati od respondenta do respondenta, se njihova storilnost ne more primerjati s storilnostjo telefonskega anketarja ali kakšnega drugega načina.

3.1.2 Telefonsko zbiranje podatkov

V anketnem raziskovanju so se začeli telefoni uporabljati že v tridesetih letih. V ZDA je v času predsedniških volitev med Roosveltom in Landnom imelo telefon 35 % celotne ameriške populacije, vendar so bili uporabniki teh pretežno republikanci. Pri napovedovanju volilnega rezultata je bila izvedena poštna anketa, kjer so bili v vzorec zbrani ljudje iz telefonskega imenika in avtomobilskega registra. Večina anketiranih je menila, da bo zmagal republikanski kandidat Landon, vendar so bili rezultati na koncu ravno obratni (Katz in Cantril v Biemer in Lyberg 2003, 193).

Telefonsko anketiranje se je množično začelo uporabljati šele okoli leta 1970 zaradi nizkih stroškov in večjega populacijskega pokritja. Groves in Khan (v Biemer in Lyberg 2003, 193) sta v svojem strokovnem članku ugotovila, da se lahko telefonske ankete po kakovosti podatkov primerjajo z osebnim terenskim anketiranjem. Danes smo zaradi nizkih stopenj odgovora zopet priča upadanju zanimanja za telefonske ankete, vendar se še vedno mnogi raziskovalci odločajo za uporabo telefonskega načina v kombinaciji z drugimi načini zbiranja podatkov.

Mobilni telefoni so se v anketiranju začeli pojavljati v devetdesetih letih. Anketiranje preko mobilnih telefonov je raziskovalcem odprlo nove možnosti in izzive pri zbiranju podatkov ter njihove kakovosti. Po Callegaru (2004, 8) poznamo več načinov anketiranja preko mobilnih telefonov. Najbolj sta razširjena načina IMP (*angl. interviewer mobile phone*) in SMP (*angl. self-administered mobile phone*).

Zaradi problemov z visokimi stopnjami neodgovora enot v anketiranju preko fiksnihi telefonov se je začelo razvijanje anketiranja preko mobilnih telefonov, ki je v osnovi isti načinu zbiranja podatkov preko fiksnega telefona, le da se pri mobilnem načinu anketira

preko mobilnega telefona. Glavna ovira mobilnega načina so visoki stroški minute klica, ki presegajo ceno klica na fiksni telefon. Poleg tega pa se zaradi prikaza številke na ekranu pojavljajo problemi neodgovora enote. Prav tako kot tudi v CATI-načinu se načrtovalci raziskav srečujejo s problemom postavitve vzorčnega okvirja zaradi neobjav mobilnih telefonskih števil v imenik, poleg tega pa smo priča poplavi mobilnih operaterjev in posledično tudi enot, ki imajo več mobilnih števil. Glavna prednost mobilnega načina zbiranja podatkov je, da imajo agencije in organizacije popolnejše pokritje ter da ta način omogoča anketiranje oseb in ne gospodinjstev, kakor je v navadi pri ostalih načinih zbiranja podatkov (Callegaro 2004, 9–11).

Telefonski način zbiranja podatkov se lahko zaradi vključenosti anketarja primerja z osebnim načinom zbiranja podatkov. V obeh metodah je prisoten učinek anketarja in socialno zaželeni odgovori, vendar je v telefonskem anketiranju stopnja učinka družbeno zaželenih odgovorov nekoliko manjša, saj anketar nima vizualnega stika z respondentom. Zagotovimo mu večjo anonimnost, zato je tudi pri občutljivih temah respondentu lažje odgovarjati, s tem pa pričakujemo višjo stopnjo odgovora kot pri terenskem osebnem načinu zbiranja podatkov. Za izvedbo telefonske ankete ne potrebujemo veliko anketarjev, poleg tega se pri telefonskem načinu lažje izvaja kontrola anketarja. Po drugi strani je telefonsko anketiranje manj fleksibilno, kajti respondentu ne moremo ponuditi vprašalnih kartic. Telefonsko anketiranje po navadi uporabljamo pri krajših (krajših od 30 minut) in manj kompleksnih raziskavah, saj respondentu preko telefona ne moramo postavljati dolgih in zahtevnih vprašanj. Prav tako smo omejeni s številom možnih odgovorov. Nepisano pravilo med raziskovalci je, da pri telefonskem anketiranju respondentu ne sme biti ponujenih več kot šest možnih odgovorov na vprašanje. Poleg tega se raziskovalci pri telefonskem načinu zbiranja podatkov srečujejo s problemom izbire skrajnih odgovorov, npr. kjer anketiranci na pet stopenjskih lestvicah izbirajo vrednosti ena ali pet, ali pa se odločijo za vrednosti »ne vem« ali »brez odgovora«. Problem telefonskega načina je, da vse osebe nimajo telefona ali ne želijo objaviti svoje številke v telefonskem imeniku, poleg tega telefonske številke niso ažurirane sproti, zato se v telefonskem raziskovanju pogosto srečujemo z neveljavnimi telefonskimi številkami ali pa tudi z napačnim kontaktom enote (Bimer in Lyberg 2003, 191–195).

V nekaterih državah uporaba telefonskega načina zbiranja podatkov preko fiksnih telefonskih priključkov kot primaren način ni možen zaradi majhne pokritosti. V številnih državah hitro

narašča število gospodinjstev, ki imajo le mobilne telefone in so po lastnostih drugačni od tistih, ki imajo fiksni telefon. Na Finskem, Češkem, v Litvi in v Latviji je danes takih že več kot 50 % gospodinjstev (Vehovar in drugi, 2010, 305). V Sloveniji je bilo v letu 2003 nekaj več kot 7 % gospodinjstev dosegljivih le preko mobilnega telefona, medtem ko delež gospodinjstev z mobilnimi telefoni presega 80 % (Vehovar in Belak 2004). V Sloveniji je po podatkih raziskave IKT 2008 delež ljudi med 18. in 50. letom, ki imajo fiksni telefon, 55 %, medtem ko je delež ljudi, ki imajo mobitel, 62,9 % (Zdešar 2008).

3.1.3 Poštno anketiranje

Pri poštnem anketiranju pošljemo anketni vprašalnik izbranemu vzorcu ljudi na domači naslov ali naslov podjetij, če imamo vzorec podjetij. Ko respondenti izpolnijo vprašalnik, ga pošljejo nazaj v raziskovalno organizacijo. Poštne ankete so eden izmed načinov samoanketiranja, zato morajo biti navodila za izpolnjevanje vprašalnika preprosta, kajti le tako se lahko izognemo veliki stopnji neodgovora enote ali spremenljivke.

Stopnja odgovora je odvisna od organizacije, ki izvaja raziskavo, od zaposlenih in njihovih sposobnosti in znanja ter v končni fazi tudi od samih lastnosti respondentov oziroma vzorčnih enot (Biemer in Lyberg 2003, 196).

Agencije se pogosto odločajo za poštno anketiranje zaradi občutno nižjih stroškov kot v načinih, kjer sodeluje anketar. Poleg tega imajo poštne ankete tudi nekatere druge prednosti pred načini, kjer sodeluje anketar (*angl. interviewer-assisted*). Tveganje za socialno zaželeno odgovore je nižje kot pri drugih načinih, poleg tega pa lahko respondentu ponudimo več možnih odgovorov kot pri telefonski anketi ali pri osebnem anketiranju (Biemer in Lyberg 2003, 196). Ena izmed največjih prednosti poštnih anket v primerjavi z ostalimi načini je ta, da lahko respondent sam odloča, kdaj bo izpolnil anketo. Tako se izognemo učinku anketarja in napakam, povzročenih s postavljanjem ali beleženjem vprašanja (Dillman 2000). Po navadi so stopnje odgovora v poštnih anketah nizke (Biemer in Lyberg 2003), zato se raziskovalci pogosto odločajo za kombinacijo še z ostalimi metodami. Poštnim anketam pogosto sledi telefonski ali terenski opomnik.

Po Biemerju in Lybergu (2003) velik problem v poštnem anketiranju predstavlja nekontrola neodgovora enote. Respondent mogoče ni dobil vprašalnika ali pa ga je izpolnil in pozabil odposlati. Lahko je vzrok tudi v raziskovalni organizaciji, če ne napišemo na vprašalnik časa, ki je na voljo za izpolnjevanje, oziroma časa odpošiljanja vprašalnika nazaj v organizacijo.

Izvedba poštnih anket je daljša od izvedbe anket z drugimi načini. Poleg tega se pri poštnih anketah soočamo s problemom nepismenosti, zato v ZDA izvajajo poštna anketa le na vzorcu populacije, ki ima pet razredov osnovne šole ali več.

3.1.4 Spletno anketiranje

Že v osemdesetih letih se je začelo anketno zbiranje podatkov preko interneta. Sprva preko elektronske pošte, v devetdesetih letih pa so začeli raziskovalci uporabljati tudi svetovni splet. Anketiranci dostopajo do njih preko brskalnikov, vprašalnik temelji na HTML-formah, odgovori pa se avtomatsko shranjujejo v bazo podatkov na strežniku. V primerjavi z anketiranjem po pošti spletno anketiranje omogoča večji razpon možnosti pri obliki vprašalnika, anketirančevo breme je manjše, avtomatično zbiranje podatkov je poenostavljeno (Lozar Manfreda in drugi 2000).

Spletno anketiranje ponuja raziskovalcem alternativo konvencionalnim metodam zbiranja podatkov, kot sta terensko ali telefonsko anketiranje, in sicer predvsem zaradi manjših stroškov in visoke hitrosti zbiranja podatkov (Lozar Manfreda in drugi, 2009). Spletne ankete lahko uporablja vsak tip informacijsko-komunikacijske tehnologije za podporo in izvedbo ankete.

Spletno anketiranje raziskovalnim organizacijam predstavlja učinkovit način, kako prihraniti stroške, saj se izognemo pošiljanju anket preko klasične pošte, kupovanju znamk, usposabljanju anketarjev ter ročnemu načinu vnašanja podatkov v bazo. Visoki so le začetni stroški, ker ta način zahteva ustrezno strojno in programsko opremo ter testiranje vprašalnikov.

Po Lozar Manfreda (2001) lahko z nižanjem stroškov v vzorec izberemo več ljudi. Z večanjem števila enot v vzorcu se večajo tudi prihranki v primerjavi z drugimi načini anketiranja, hkrati pa se z večanjem vzorca zmanjšujejo vzorčne napake.

Poleg nizkih stroškov je ena izmed glavnih prednosti spletnega anketiranja čas zbiranja podatkov (Lozar Manfreda in drugi 2000). Podatke lahko v primerjavi s poštnimi anketami zberemo v nekaj dneh.

Računalniško zbiranje podatkov oziroma sprotno beleženje odgovorov enot v bazo na strežniku omogoča raziskovalcem hitrejšo obdelavo podatkov v končni fazi raziskovanja. Danes je na spletu mogoče najti številna orodja, ki že z malo računalniškega znanja omogočajo posamezniku izvedbo spletne ankete in tudi izvedbo deskriptivnih statistik na spletu. Spletno anketiranje pa se že nekaj časa sooča z dvema velikima problemoma, problemom nepokritja in neodgovora.

Po poročilu Evropske komisije (2009) se je evropsko povprečje redne uporabe interneta od leta 2005 do 2008 povečalo za 13 odstotnih točk (iz 43 % na 56 %). Slovenija se po rednih uporabnikih interneta uvršča malo pod povprečje EU27. Pokritost s DSL-tehnologijo v Sloveniji je bil leta 2008 po podatkih raziskave »Raba interneta v Sloveniji« 92-odstotna, medtem ko je bila leta 2005 55-odstotna. Kljub velikemu skoku pri pokritju pa ostaja 8 % populacije nepokrite, zato spletno anketiranje v uradnih anketah ne pride v poštev. Največjo pokritost z DSL-tehnologijo imajo Velika Britanija, Belgija, Danska, Luksemburg in Francija, medtem ko imajo najnižjo pokritost Poljska, Grčija, Bolgarija ter Romunija (68 %). Največji problem spletnih anket predstavljajo neodgovori enot. Stopnje odgovora so v spletnem anketiranju najnižje od vseh načinov. Spletno anketiranje se po navadi uporablja v internetnih panelnih raziskavah, kjer je stopnja odgovora bistveno večja. Anketiranje po internetu bo postalo enakovredno drugim metodam, ko bo dokazano, da njegove anketne napake (napake zaradi neodgovorov, nepokritja, vzorčenja in samega merjenja) niso večje od tistih, ki so dobljene z drugimi metodami (Lozar Manfreda in drugi 2000).

3.2 Izbira ustrezne metode

Izbira ustrezne metode zbiranja podatkov je odvisna od številnih faktorjev, ki naredijo posamezno metodo bolj ali manj primerno (Biemer in Lyberg 2003, 205). Metodo zbiranja podatkov lahko določimo na podlagi treh dimenzij: raziskovane populacije, razpoložljivih virov ter značilnosti vprašalnika.

3.2.1 Odvisnost od raziskovane populacije

Kot prvo se odločimo za izbiro metode na podlagi populacije, ki jo želimo anketirati. Vsaka metoda omogoča doseganje določene populacije, poleg tega pa lahko izbiramo populacijo tudi na podlagi raziskovane tematike. Za raziskave, ki temeljijo na nacionalno reprezentativnem vzorcu, je najboljša metoda terensko anketiranje, ker imajo državni statistični uradi, ki vodijo registre prebivalstva (v Sloveniji Centralni register prebivalstva – CRP), dostop do vseh domačih naslovov prebivalcev države. V teoriji lahko anketar vsakega državljanca obišče na domu, vendar moramo v praksi upoštevati družbene in geografske specifične posamezne države. Prav tako kot za terensko anketiranje se zaradi dostopnosti podatkov državljanov lahko odločimo tudi za pošto anketiranje, medtem ko načini, ki temeljijo na samoanketiranju, zaradi faktorja pismenosti ne pridejo vedno v poštev. Za metodo raziskovanja na nacionalno reprezentativnem vzorcu lahko izberemo tudi telefonsko anketiranje. V Sloveniji SURS vzorči iz telefonskega imenika števil (angl. *list based approach*), medtem ko je v ZDA in Angliji razširjen sistem RDD (angl. *random digit dialing*). Zaradi opuščanja fiksnih telefonskih priključkov in uporabe le mobilnih telefonov je sistem vzorčenja preko telefonskega imenika števil čedalje manj primeren, saj tvegamo napako pokritosti. Spletno anketiranje je primeren način zbiranja podatkov za specifične populacije, kot so študenti, zaposleni in podjetja, za katera vemo, da imajo internetni dostop. Na Nizozemskem in ZDA so začeli uporabljati nacionalno reprezentativne internetne panele, ki temeljijo na naključnem vzorčenju. Internetni paneli postajajo vse bolj popularni v sodobnem raziskovanju (ESOMAR 2005).

3.2.2 Odvisnost od virov

Metodo zbiranja podatkov lahko izberemo tudi na podlagi administrativnih in stroškovnih omejitev (Czaja in Blair 2005). Te omejitve obsegajo čas implementacije in izvedbo raziskave. Vsak način ima svoje specifične stroškovne lastnosti. Terensko anketiranje velja za najdražje zaradi potovalnih stroškov. Tudi telefonsko anketiranje postaja vse dražje, saj morajo anketarji z latentnim padcem sodelovanja opraviti več poskusov kontakta, hkrati pa narašča strošek dela anketarja. Groves in Kahn (v Gaziano 2005, 128) sta leta 1979 v svoji primerjalni študiji zaključila, da je telefonsko anketiranje za več kot polovico cenejše od osebnega anketiranja. Poštne ankete so še cenejše od telefonskih anket, vendar so nižji stroški povezani z daljšo izvedbo ankete. Anketiranci morajo namreč izpolnjene vprašalnike poslati nazaj organizaciji. Čas izvedbe poštne ankete podaljšujejo tudi poštni opomniki (Dillman 2000). Razvoj interneta kot anketnega orodja je pospešil proces zbiranja podatkov, vendar so skladno s hitrostjo zbiranja podatkov povezani tudi začetni stroški nakupa računalniške opreme in programiranja. Takoj ko torej poznamo začetne stroške, lahko ocenimo, kateri način zbiranja podatkov je za nas najbolj sprejemljiv. Hkrati moramo biti zaradi reprezentativnosti pozorni na ustrezno pokritost in velikost vzorca.

3.2.3 Odvisnost od vprašalnika

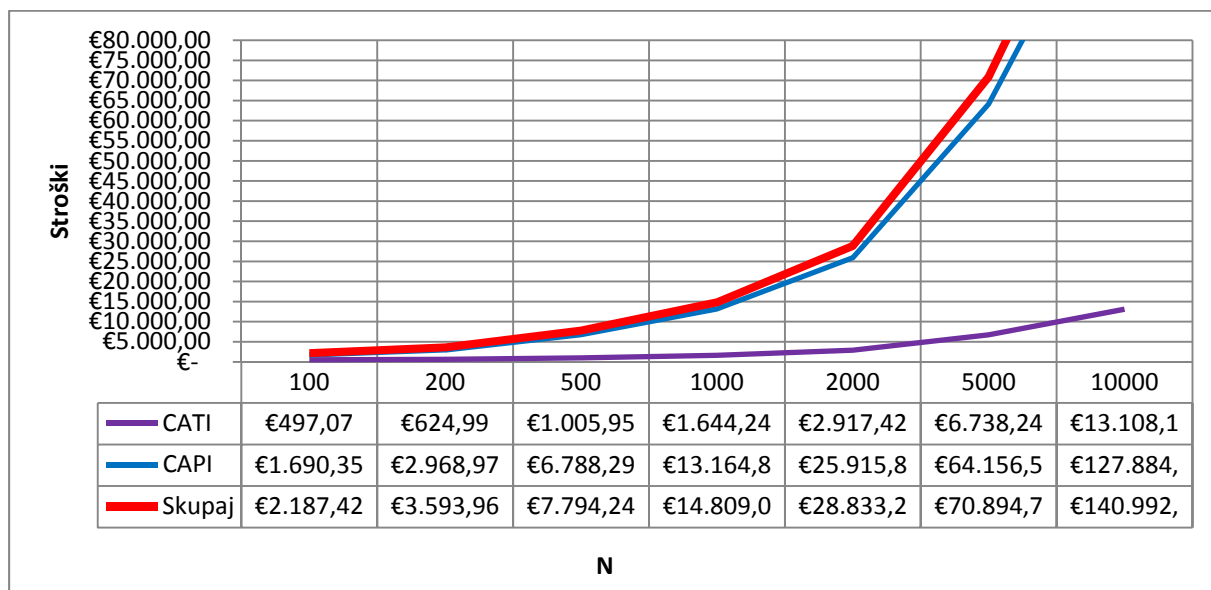
Kot tretje lahko metodo zbiranja podatkov izberemo na podlagi zahtevnosti vprašalnika (Cazaja in Blair 2005). Vprašalniki se razlikujejo po dolžini in kompleksnosti. V primeru, da imamo kompleksen in dolg vprašalnik, se po navadi ne odločimo za način samoanketiranja, ker bi s tem tvegali prevelike napake pri odgovarjanju na vprašanja in prenehanju med njimi (npr. preskakovanju pogojnih vprašanj). Razvoj računalniško podprtega anketiranja je omogočil uporabo vse kompleksnejših vprašalnikov. Poleg tega so lahko spletne ankete programirane tako, da nas računalnik vpraša na podlagi odgovora prejšnjega vprašanja. Če imamo daljše vprašalnike, se po navadi odločimo za terensko ali telefonsko anketiranje, kjer lahko na željo anketiranca anketiranje tudi prekinemo in ga nadaljujemo kasneje. Rezultati empiričnih raziskav o vplivu na kakovost podatkov pri dolgih vprašalnikih se zelo razlikujejo (Dillman 2000). Več je empiričnih dokazov, ki nakazujejo pomembnost krajšega vprašalnika pri načinih samoanketiranja. Pri spletnih anketah je še toliko višja verjetnost, da anketiranec

neha odgovarjati na vprašanja zaradi ostalih dejavnikov, povezanih z uporabo interneta in odprtih aplikaciji (Groves in drugi 2004).

3.3 Vidik stroškov

Stroški se med načini zbiranja podatkov zelo razlikujejo med seboj. Na splošno lahko rečemo, da je najdražji način zbiranja podatkov terensko anketiranje, sledi mu telefonsko, pošto in spletno, ki je najcenejše. Za lažjo predstavo o razmerjih stroškov med različnimi načini smo se odločil prikazati razmerje med terenskim (CAPI) in telefonskim (CATI) načinom, in sicer z upoštevanjem stopnje odgovora posameznega načina v raziskavi IKT_GOS 2008 (Belak in Zdešar 2008, 2). Pri terenskem načinu je bila stopnja odgovora 76,90 %, medtem ko je bila pri telefonskem 40,90 %. Sedem različnih velikosti vzorcev bomo upoštevali kot začetno velikost raziskovanega vzorca. Stroške smo določili na podlagi ocenitve stroškov in parametrov posameznega načina (glej Prilogo A, tabeli A.1 in A.2).

Slika 3.1: Simulacija stroškov in velikosti vzorca



V nadaljevanju smo naredili isto simulacijo v primeru 100-odstotne udeležbe respondentov v obeh načinih in ob predpostavki, da je bilo preko telefona anketiranih 50 % preko fiksnega telefona in 50 % preko mobilnega.

Slika 3.2: Simulacija stroškov glede na način in velikost vzorca ob predpostavki o 100-odstotni stopnji sodelovanja

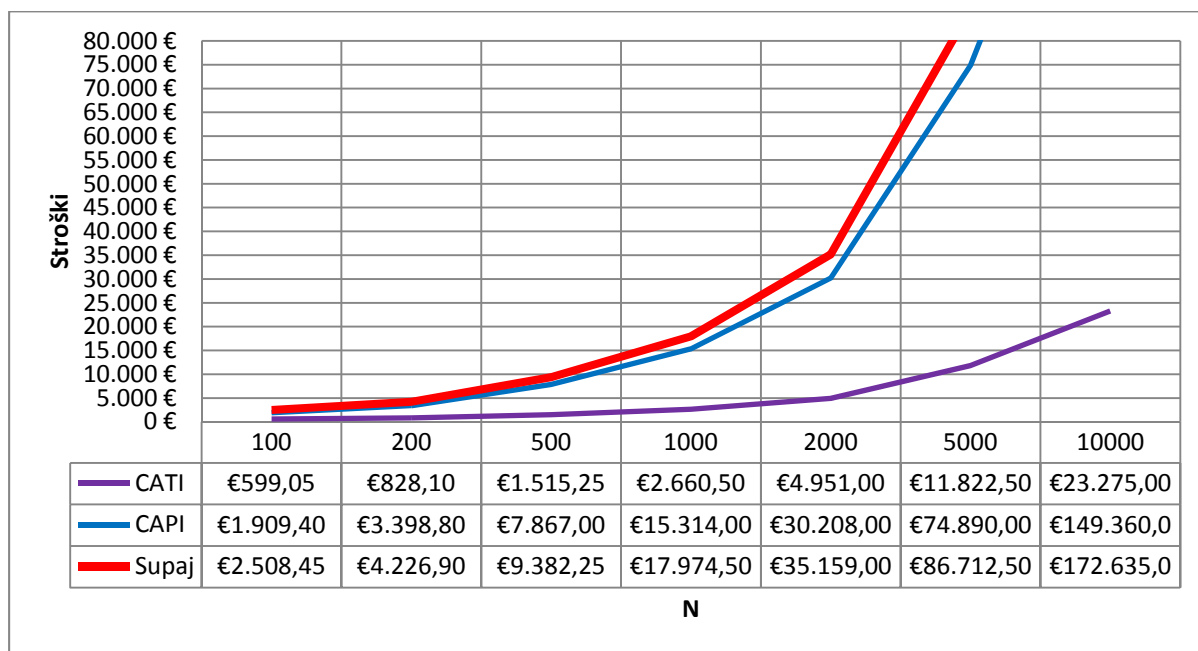


Tabela 3.2: Razmerja stroškov med CATI- in CAPI-načinom anketiranja

Stopnje IKT	3,401	4,751	6,748	8,001	8,883	9,521	9,756
100-odstotna stopnja sodelovanja	3,187	4,104	5,191	5,756	6,101	6,334	6,417

Iz slik opazimo, da je terensko anketiranje dražje od telefonskega anketiranja. V primeru velikosti vzorca $n = 500$ in upoštevanju stopnje odgovora posameznega načina iz raziskave IKT_GOS 2008 je terenski način zbiranja podatkov 6,75-krat dražji od telefonskega načina. Z večanjem vzorca se razlike v stroških le še povečujejo, tako je terenski način pri $n = 5000$ 9,52-krat, pri $n = 10000$ pa 9,77-krat dražji od telefonskega. Pri upoštevanju 100-odstotne stopnje sodelovanja v obeh načinih se razlike nekoliko zmanjšajo. V primerjavi s prejšnjo ocenitvijo so razlike manjše, kajti stopnja odgovora je pri obeh načinih enaka, zato so tudi razlike med načinoma manjše. Pri velikosti vzorca $n = 500$ je 5,2-krat, pri $n = 10000$ pa je 6,4-krat dražji terenski način od telefonskega. Iz Tabele 3.2 dobimo koristno informacijo o razmerju stroškov in velikosti vzorca. Največje razlike med načini so pri majhnih vzorcih ($n = 100, 200, 500$). Z večanjem vzorca se stroški sicer povečujejo, vendar se razmerje med stroški ustali. Pri majhnih vzorcih je med velikostjo vzorca 100 in 500 terenski način še enkrat dražji, medtem ko je pri velikih vzorcih (med 1000 in 10000) terenski način za 0,22-krat dražji od telefonskega načina.

4 Kombiniranje načinov anketiranja

Skladno s številom ponujenih metod zbiranja podatkov se porajajo tudi vprašanja o tem, katero metodo izbrati glede na stopnjo odgovora, mersko napako in glede na razpoložljive vire. Raziskovalci naj bi se pri oblikovanju načrta raziskave odločali o izbiri načina zbiranja podatkov na podlagi kakovosti podatkov (de Leeuw 2005), vendar se danes pri izbiri metode zbiranja podatkov raziskovalci odločajo predvsem na podlagi administrativnih in stroškovnih virov (Czaja in Blair 2005). Vedno več raziskovalcev se odloča za kombinacijo različnih načinov anketiranja, saj lahko z optimalno kombinacijo načinov dobimo bolj kakovostne podatke ob nižjih stroških kot pri raziskavah, ki uporabljajo le en način, bodisi v vseh fazah raziskovalnega procesa ali le v fazi zbiranja podatkov (de Leeuw 2005).

4.1 Razlogi za uporabo kombiniranih načinov

Glavni razlog uporabe kombiniranih načinov anketiranja v raziskovanju je nadomestitev slabosti enega načina s prednostmi drugega načina in s tem izboljšanje kakovosti pridobljenih podatkov.

Za kombinirane načine zbiranja podatkov se odločamo zaradi treh glavnih razlogov (Jowell 1998, 174):

1. poveča se stopnja odgovora in sodelovanja, zmanjšajo se napake,
2. zmanjšajo se stroški raziskave,
3. hiter razvoja IKT ter potencialnih prehodov na nove načine zbiranja podatkov.

V modernem raziskovanju smo priča generalnemu upadanju stopnje odgovora (*angl. response rate*) v vseh načinih zbiranja podatkov. Stopnja neodgovora se večja zaradi nekontakta in zavrnitev sodelovanja. Stopnja nekontakta se povečuje predvsem pri uporabi tradicionalnih načinov zbiranja podatkov (Couper 2002). Vse več ljudi živi v t. i. zaprtih skupnosti ali uporablja mobilni ali fiksni telefon le za lastne potrebe. Z uporabo posameznega načina anketiranja lahko dostopamo do posameznih podpopulacij in tako manjšamo napako pokritosti. Zaradi splošnega prepričanja in upadanja ugleda raziskovanja se je povečala tudi stopnja zavrnitve sodelovanja. Istočasno pa so se povečali stroški, deloma zaradi poskusov ublažitve napake pokritosti in napake neodgovora; daljša se čas zbiranja podatkov (anketarji

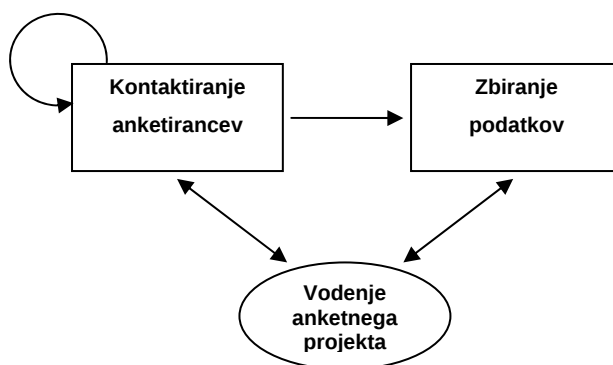
morajo večkrat klicati ali obiskovati ljudi, preden vzpostavijo kontakt z izbrano osebo), poleg tega se večja prepričanje ljudi, da morajo biti za sodelovanje v raziskavi nagrajeni. Zaradi vse močnejših vplivov stroškovnih in administrativnih omejitev na raziskovalni proces smo danes priča sistematičnemu opuščanju raziskav, ki temeljijo le na enem načinu zbiranja podatkov (Dillman 2000).

4.2 Klasifikacija kombiniranih načinov

Raziskovalni proces lahko razdelimo na dve glavni fazi:

- faza kontaktiranja anketirancev (začetno kontaktiranje in opomniki za nerespondente)
- faza zbiranja podatkov

Slika 4.1: Ločevanje faze kontaktiranja in zbiranja podatkov



Vir: Vehovar in Lozar Manfreda (2008, 186).

Za razumevanje širokih možnosti uporabe kombiniranih načinov anketiranja in velikega števila možnih implementacij je pomembno razlikovanje med posameznimi fazami v raziskovalnem procesu (de Leeuw 2005). V vsaki fazi raziskovalnega procesa lahko uporabimo drugačen način. V fazi najave ankete lahko anketirancem na primer pošljemo pismo, za termin anketiranja se lahko dogovorimo preko telefona, medtem ko je glavni način zbiranja podatkov izveden na terenu. Po de Leeuw (2005) načine kombiniranja metod v vseh fazah raziskovalnega procesa imenujemo sistem kombiniranih načinov (*angl. mixed-mode survey system*).

Dillman (2000, 219–222) razlikuje med petimi različnimi načini kombiniranja:

1. uporaba različnih načinov zbiranja podatkov za različne enote vzorca,
2. uporaba različnih načinov zbiranja podatkov v posameznih korakih panelne raziskave,
3. uporaba različnih načinov za posamezne dele vprašalnika,
4. uporaba različnih načinov za zbiranje podatkov iz različnih populacij (mednarodne raziskave),
5. uporaba kombiniranih načinov za komunikacijo med anketarji in respondenti v različnih fazah raziskovalnega procesa.

Edith de Leeuw (2005) je v svojem strokovnem članku dopolnila Dillmanovo in Baldnovo klasifikacijo kombiniranja načinov. Kombinirane načine anketiranja klasificira glede na posamezno fazo v raziskovalnem procesu, in sicer fazo kontaktiranja, fazo odgovora in fazo opominjanja za nerespondente.

V nadaljevanju bomo predstavili razširjeno klasifikacijo kombiniranja načinov po de Leeuw (2005).

4.2.1 Kombiniranje načinov v fazi začetnega kontaktiranja

4.2.1.1 Način najave ankete je različen od načina zbiranja podatkov

Ta način uporabljajo raziskovalci za povečanje legitimnosti in za povečanje zaupanja v raziskavo. Za ta način se raziskovalci odločajo predvsem v raziskavah, kjer glavni način zbiranja podatkov respondentu ne daje občutka legitimnosti raziskave (telefonsko in spletno anketiranje). De Leeuw, Hox, Korendijk, Lansvelt- Mulders in Callegaro (v de Leeuw 2005, 237) so empirično dokazali, da poštna najava telefonske ankete zviša stopnjo odgovora. Prav tako so Paxson, Dillman in Tarnai (v de Leeuw 2005, 239) dokazali, da telefonska najava poštne ali spletne ankete poveča stopnjo odgovora. Najpogosteje uporabljeni pristopi so poštna najava telefonske ali spletne ankete in telefonska ali poštna najava terenske ankete. Prav tako se lahko odločimo za najavo ankete, kjer anketiranec sam izbere način anketiranja, npr. kjer anketiranec dobi pismo z najavo ankete, sam pa izpolni vprašalnik na spletu ali pa se

odloči za poštno anketo (Lozar Manfreda, Vehovar in Batagelj 2001). Pri anketiranju podjetij se raziskovalci najpogosteje odločijo za telefonsko najavo poštne ankete.

4.2.1.2 Način vabila, presejalnega (angl. screening) vprašalnika ali izbora enot je različen od načina zbiranja podatkov

Podobno kot za najavo ankete je mogoče drugačen način kot za samo zbiranje podatkov uporabiti tudi za neposredno vabilo k sodelovanju. Ta način je nujen pri spletnih anketah splošne populacije (razen v panelih). Pogosto se namreč zgodi, da nimamo elektronskih naslovov anketirancev, ta pristop pa lahko uporabljamo tudi pri zmanjševanju napak nepokritja v drugih načinih. V telefonski anketi lahko na primer posameznike, ki jim nismo mogli določiti telefonske številke, s poštnim vabilom povabimo, da za anketiranje pokličejo brezplačno telefonsko številko. Z razvojem spletnega anketiranja postajajo vse pogostejši kombinirani načini, pri katerih v prvi fazi izvedemo presejalni vprašalnik oziroma izbor enot. V raziskavah, kjer je glavni način ali en izmed načinov zbiranja podatkov spletna anketa, najprej s telefonskim ali poštnim presejalnim vprašalnikom določimo enote, ki imajo splet, in nato pridobimo kontakte za vabila po elektronski pošti. Najpogosteje se zaradi omejenega proračuna raziskave izvede presejalni vprašalnik preko telefona (de Leeuw 2005).

4.2.2 Kombiniranje načinov v fazi zbiranja podatkov

4.2.2.1 Različni načini za različne enote istega vzorca v eni časovni točki in z enakim vprašalnikom

Uporaba različnih načinov za različne enote v vzorcu postaja vedno pogostejša metoda v anketnem raziskovanju. Ločimo tri načine tega pristopa, ki so opisani v nadaljevanju.

4.2.2.1.1 Izbirno

Anketiranec med ponujenimi načini anketiranja sam izbere sebi najprimernejšega. Najpogosteje se ta metoda uporablja pri načinih, kjer ena izmed metod ni dostopna vsem

anketirancem. Tako na primer anketirancu poleg vabila za sodelovanje v spletni anketi posredujemo še poštni vprašalnik ali pa ga prosimo, naj pokliče na brezplačno telefonsko številko za telefonsko anketiranje. Ljudje so pri tej metodi v zadregi med izbiro načina, zato velikokrat pride do neodgovora enote.

4.2.2.1.2 Sočasno

Raziskovalec sam določi, kateri način je najprimernejši za posameznega anketiranca. Raziskovalci se odločajo na podlagi kontaktnih informacij izbranih oseb. Primer take raziskave je IKT_GOS 2008, kjer so osebe, ki so jim določili telefonske številke, anketirali preko telefona, osebe, katerim niso mogli določiti telefonskih števil, pa na terenu. Odločitev o načinu anketiranja lahko temelji tudi na predhodnem poznavanju značilnosti določenih delov populacije (podpopulacije), kjer se odločimo za način, ki je za posamezno podpopulacijo najoptimalnejši.

4.2.2.1.3 Zaporedno

V zaporednem načinu so uporabljeni različni načini anketiranja v različnih fazah anketiranja. Ta način je učinkovit pri zmanjševanju stopnje neodgovora in je skladen z načeli Dillmanove »po meri oblikovane metode anketiranja« (angl. Tailored Design Method – TDM), ki temelji na teoriji družbene menjave. TDM predstavljajo postopki, kako ob manjših stroških ustvariti zaupanje anketirancev v raziskavo (Dillman 2000). V tem primeru nerespondente anketiramo z drugim načinom. Najpogosteje se raziskovalci odločijo za zaporedje, ko se v primeru kombiniranih načinov zbiranja podatkov odločijo najprej za cenejši način (npr. spletno anketo, nerespondente poskušamo anketirati preko dražjega telefonskega načina, nerespondente na telefonski način pa potem poskušamo anketirati z najdražjim terenskim načinom). S tem pristopom neposredno vplivamo na stroške in stopnjo sodelovanja (Voogt in Saris 2005). Poleg zniževanja stopnje neodgovora pa tako zmanjšujemo tudi napako nepokritosti; če npr. starejši anketiranec ne more odgovoriti na anketo preko svetovnega spleta, lahko njegove podatke zberemo preko telefonske ali terenske ankete. Ta način je delno uporabljen v raziskavi IKT_GOS 2008, kjer so telefonske nerespondente poskušali anketirati

še na terenu, pri čemer so bili izjemno uspešni. Stopnja odgovora v drugi fazi je bila 58,4-odstotna (Belak in Zdešar 2008, 17).

Uporaba različnih načinov anketiranja za različne enote iz vzorca lahko pomembno prispeva k večji stroškovni optimalnosti anketiranja ter višji stopnji sodelovanja in pokritja. Hkrati pa takšen pristop predstavlja potencialno grožnjo kakovosti podatkov zaradi možnosti nastanka učinkov načina anketiranja (napak, povezni s posameznim načinom) in s tem povečevanja merske napake.

4.2.2.2 Različni načini za različne dele vprašalnika za vse enote iz istega vzorca v eni časovni točki in z enakim vprašalnikom

Omenjena metoda se največkrat uporablja pri zmanjševanju stopnje družbeno zaželenih odgovorov ali zavračanju odgovarjanja na vprašanja. Pri terenskem anketiranju je občutljiva vprašanja mogoče izvesti z načinom samoanketiranja (npr. z metodo CAPI), enako kot pri Anketi o žrtvah kriminala – AŽK (Belak in Kontelj 2008), kjer je anketar anketirancu ponudil odgovarjanje na vprašanja preko prenosnega računalnika.

4.2.2.3 Različni načini v različnih časovnih točkah na enotah v istem vzorcu (panel)

Ta oblika kombiniranja načinov anketiranja se uporablja predvsem v panelnih raziskavah, kjer isti anketiranec sodeluje večkrat, vendar je v različnih valovih raziskave vključen v različne načine anketiranja. Najpogosteje se raziskovalci v prvem valu odločijo uporabiti terensko anketiranje, da se izpostavi legitimnost raziskave, v naslednjem valu pa iste anketirance anketiramo s cenejšimi načini (de Leeuw 2005).

Z razvojem interneta in IKT so se začeli uporabljati tudi internetni paneli, ki nastanejo po predhodnem valu, opravljenim s telefonsko ali terensko anketo, kasnejši valovi pa potekajo izključno s spletnimi anketami. Taki paneli so bistveno cenejši od klasičnih panelov.

Menjavanje načinov med valovi anketiranja je problematično predvsem zaradi mešanja dejanskih sprememb v času z razlikami, ki so posledica uporabljenega načina anketiranja.

4.2.2.4 Različni načini v različnih časovnih točkah na različnih vzorcih (enotah)

V tem primeru gre za raziskave, ki se izvajajo v več valovih (npr. mednarodne raziskave ESS) in v vsakem izmed njih rekrutiramo nov vzorec ljudi, kjer lahko na določeni točki zamenjamo način anketiranja. V vseh valovih je lahko uporabljen enak ali drugačen vprašalnik, kar je odvisno od raziskave. V primeru, da v posameznih valovih uporabljamo različne načine anketiranja, gre za posebno obliko kombiniranja načinov anketiranja. Razlogi za menjavo načinov so običajno posledica stroškovnih omejitev, ki zahtevajo prehod na cenejše in bolj praktične načine anketiranja.

Podobno kot v zgornjem primeru velja, da menjava načina anketiranja lahko zabriše vzroke nastanka sprememb med opazovanji, ki so lahko posledica merske napake zaradi načina anketiranja in ne dejanskih sprememb.

4.2.3 Kombiniranje načinov v fazi opomnikov

Za povečevanje stopnje odgovora se po navadi raziskovalci odločajo za uporabo opomnikov, ki so lahko posredovani na isti način kot glavna metoda zbiranja podatkov (npr. poštni opomnik za poštno anketo ali telefonski opomnik za telefonsko anketo) ali na različen način (npr. telefonski opomnik za spletno anketo).

Uporaba različnih načinov opomnikov poveča stopnjo odgovora, saj dobi anketiranec informacijo o raziskavi preko različnih virov. Pogosto pa se raziskovalci odločajo za menjavo načinov opomnikov v posameznih fazah njihovega posredovanja (npr. najprej poštni, nato pa še telefonski opomnik za spletno anketo).

Sam način opomnikov, brez menjave načina anketiranja v kasnejših fazah, ne povzroča učinkov načina anketiranja v ožjem smislu, lahko pa vpliva na specifične vzorce neodgovora (npr. določen način opomnika bolj poviša stopnjo sodelovanja specifičnih skupin).

4.2.4 Druge oblike kombiniranja načinov

Zgoraj smo opisali najbolj pogoste načine kombiniranja načinov, vendar obstajajo še drugi specifični načini, in sicer:

- sočasna izvedba anketiranja na enotah iz različnih vzorčnih okvirov (npr. kombinacija fiksnih telefonskih števil iz telefonskega imenika in RDD-mobilnih števil),
- različni načini najave za različne enote iz vzorca,
- uporaba različnih načinov presejalnega vprašalnika v fazi začetnega kontaktiranja, ki je lahko izbirna, sočasna ali zaporedna (podobno kot kombiniranje v fazi glavnega zbiranja podatkov),
- različni načini opomnikov za različne enote iz vzorca,
- drugi načini.

4.3 Dileme kombiniranja načinov

Raziskovalci imajo pri načrtovanju raziskav veliko možnosti uporabe kombiniranih načinov v posameznih fazah raziskovalnega procesa. Priporočil o uporabi posamezne kombinacije metod zaradi specifikacije posameznega raziskovalnega problema praktično ni. Kljub temu lahko na podlagi nekaterih empiričnih raziskav postavimo kombiniranim načinom osnove, na podlagi katerih se raziskovalci lažje odločijo o uporabi različnih načinov v posameznem delu raziskovalnega procesa. Če se odločimo za kombiniranje načinov samo v fazah kontaktiranja, presejalnega vprašalnika in opomnika, medtem ko za glavno metodo zbiranja podatkov uporabljamo samo en način, so posledice uporabe kombiniranih načinov na kakovost podatkov le pozitivne (de Leeuw 2005). V primeru, da se odločimo za kombiniranje načinov v fazi zbiranja podatkov, pa so posledice zaradi dilem o združevanju zbranih podatkih po posameznem načinu nejasne. Le v primeru uporabe bolj anonimnih načinov anektiranja v primeru druge metode zbiranja podatkov (npr. terensko anketiranje, za občutljiva vprašanja, kjer tvegamo družbeno želene odgovore, pa anketiranec odgovori na njih preko računalnika) so bili z empiričnimi raziskavami do sedaj dokazani pozitivni učinki na kakovost podatkov (de Leeuw 2005, 247). V vseh ostalih primerih kombiniranja načinov so lahko učinki na

kakovost podatkov negativni. Če je na primer presejalni vprašalnik izveden z drugim načinom kot glavni anketni vprašalnik, se lahko pojavi učinek posameznega načina, kar vpliva na kakovost podatkov. S kombiniranjem različnih načinov anketiranja ima raziskovalec možnost kontroliranja stopnje posameznega učinka načinov, vendar se ob nenačrtovani in neargumentirani uporabi različnih načinov lahko stopnja učinka načina celo poveča, zato moramo biti pri načrtovanju uporabe kombiniranih načinov še posebej previdni (Dillman 2000).

Raziskovalci se morajo pri načrtovanju raziskave, kjer bodo uporabljali kombinirane načine, odločiti, katero metodo izbrati glede na minimizacijo učinkov načina. De Leuw (2005) opozarja, da morajo raziskovalci v fazi načrtovanja raziskave razlikovati med dvema povsem različnima modeloma kombiniranih načinov:

- modelom, kjer uporabljamo en glavni način zbiranja podatkov, medtem ko je drugi način prvemu dopolnilni (npr. uporabljamo drugi način za neodgovore ali v panelnih raziskavah, kjer v prvem valu naredimo terensko rekrutacijo anketirancev, v drugih valih pa uporabljamo telefonski način),
- modelom, kjer uporabimo več enakovrednih načinov zbiranja podatkov hkrati (npr. anketirancu ponudimo možnost telefonske, spletne ali poštne ankete).

V primeru, ko se odločimo le za eno glavno metodo zbiranja podatkov, mora biti ta izbrana tako, da v čim večji meri izkoristi svoje pozitivne značilnosti, nadomestne metode pa izberemo tako, da nadomestimo slabosti prve metode (de Leeuw 2005, 248). Dillman (2000) v svojem delu predstavi načela, kako v raziskavah, ki uporabljajo poleg glavnega načina zbiranja podatkov še nadomestne načine, zmanjšati razlike med glavnim in nadomestnim načinom. Po njegovem je glavni vzrok učinkov načina drugače oblikovan vprašalnik (npr. terenski vprašalnik se razlikuje od spletnega). Problem vprašalnikov je, da se danes sestavljajo glede na posamezen način, namesto da bi se vsi prilagodili v tako obliko, da bi ga lahko uporabljali v vseh načinih, vendar je zaradi razlik med komunikacijskimi stimulatorji (npr. stimulatorji pri slušnih načinih se razlikujejo od tistih v vidnih) poenotenje vprašalnika še vedno velik izziv (Dillman 2000, 224). V panelnih raziskavah, kjer se raziskovalci odločajo za kombinirane načine (npr. terenska anketa v prvem valu ter telefonska za vse ostale vale), se

srečujejo s še večjim problemom; težko je namreč ločiti med učinkom načina in časovnim učinkom, ki povzročata razlike med odgovori. V mednarodnem raziskovanju (npr. European Social Survey – ESS) se prav tako srečujemo s problemom kombiniranja načinov. V nekaterih državah se izvajajo raziskave z uporabo enega načina, medtem ko v drugih z uporabo dveh načinov. V državah, ki uporabljajo dva načina (npr. poštno in telefonsko), se zato ne vzame v raziskovalni vzorec ljudi iz celotne populacije, temveč le enote iz ene podpopulacije (npr. v vzorec vzamemo enote s telefonskim dostopom, te razdelimo na tiste, ki bodo anketirani poštno, in tiste, ki bodo anketirani preko telefona).

4.4 Klasifikacija anket SURS

Nacionalni statistični urad Republike Slovenije (SURS) je glavni izvajalec in povezovalec dela na področju državne statistike. Poleg povezovanja in usklajevanja statističnega sistema sodijo med njegove najpomembnejše naloge še mednarodno sodelovanje, določanje metodoloških in klasifikacijskih standardov, predvidevanje potreb uporabnikov, zbiranje, obdelava in izkazovanje podatkov ter skrb za njihovo zaupnost (SURS 2011, 20. avgust). SURS izvaja raziskave gospodinjstev in podjetij na področjih demografije, naravoslovja, ekonomije in drugo. Spada v sklop vladnih služb, kar pomeni, da se vlada na podlagi statističnih podatkov odloča o spremembah politike posameznega področja, zato je glavna naloga SURS-a zagotovitev kakovostnih podatkov, hkrati pa se financira iz davkoplačevalskega denarja, zato je optimizacija stroškov in s tem tudi uporaba kombiniranih načinov še toliko pomembnejša.

Za pomoč pri klasifikaciji raziskav gospodinjstev ter podjetij smo uporabili razširjeno klasifikacijo kombiniranih načinov po de Leeuw (2005). Analizirali smo metodologije osmih najnovejših raziskav gospodinjstev in dveh raziskav podjetij (IKT_PODJ in PA-GRAD/M). Raziskave smo izbrali glede na tri vidike: dostopnost metodologije raziskave, način zbiranja podatkov in tip raziskave.

Tabela 4.1: Klasifikacija načinov anketiranja po de Leeuw (2005)

Raziskava		Ankete gospodinjstva								Ankete podjetji	
		IKT_GOS	AMP	TU-ČAP	SPK-AZK	AIO	APEVG	KDZ	SILC	IKT_PODJ	PA-GRAD/M
Faza začetnega kontakta.	Način najave ankete različen od načina anketiranja	DA	DA	DA	DA	DA	DA	delno	DA		
	Način vabljenja, presejalnega vprašalnika ali izbora enot različen od načina zbiranja podatkov	/	/	/	/	/	/	/	/	/	/
Faza zbiranja podatkov	Različni načini za različne enote iz enega vzorca v eni časovni točki in z enakim vprašalnikom	DA			DA	DA	DA	DA	DA		(DA) Za manjkajoče odg.
	Različni načini za različne dele vprašalnika za vse enote iz enega vzorca v eni časovni točki in z enakim vprašalnikom										
	Različni načini v različnih časovnih točkah na istem vzorcu										
	Različni načini v različnih časovnih točkah na različnih vzorcih		/	/	/	/	/	/	/	/	/
Faza opomnikov	Način opomnikov različen od načina zbiranja podatkov		/	/	/	/	/	/	/	/	/

Legenda: / – ni podatka, IKT_GOS – Anketa o uporabi IKT v gospodinjstvih in posameznikih, AMP – Anketa o mnenju potrošnikov, TU-ČAP – Turistična potovanja domačega prebivalstva, SPK-AZK – Anketa o žrtvah kriminala, AIO – Anketa o izobraževanju odraslih, APEVG – Anketa o porabi energije in goriv v gospodinjstvih, KDZ – Kariera doktorjev znanosti, SILC – Raziskava o življenjskih pogojih, IKT_PODJ – Uporaba IKT v podjetjih, PA-GRAD/M – Anketa o poslovnih tendencah v gradbeništvu

Tabela 4.2 : Klasifikacija raziskav SURS po dopolnitvi klasifikacije Edith de Leeuw (2005)

Način	Opis	Ankete gospodinjstva								Ankete podjetji	
		IKT_GOS	AMP	TU-ČAP	SPK-AZK	AIO	APEVG	KDZ	SILC	IKT_PO DJ	PA- GRAD/M
4.2.1.	Kombiniranje načinov v fazi začetnega kontaktiranja	DA	DA	DA	DA	DA	DA	DA	DA		
4.2.1.1.	Način najave ankete je različen od načina zbiranja podatkov	DA	DA	DA	DA	DA	DA	DA	DA		
4.2.1.2.	Način vabljenja, presejalnega (screening) vprašalnika ali izbora enot je različen od načina zbiranja podatkov										
4.2.2.	Kombiniranje načinov v fazi zbiranja podatkov	DA			DA	DA	DA	DA			
4.2.2.1.	Različni načini za različne enote iz istega vzorca v eni časovni točki in z enakim vprašalnikom	DA			DA	DA	DA	DA			DA
4.2.2.1.1.	Izbirno				delno						
4.2.2.1.2.	Sočasno	DA			DA	DA	DA	DA			
4.2.2.1.3.	Zaporedno	delno			delno	delno					DA
4.2.2.2.	Različni načini za različne dele vprašalnika za vse enote iz istega vzorca v eni časovni točki in z enakim vprašalnikom										
4.2.2.3.	Različni načini v različnih časovnih točkah na enotah istega vzorca (panel)										
4.2.2.4.	Različni načini v različnih časovnih točkah na različnih vzorcih (enotah)								DA		
4.2.3.	Kombiniranje načinov v fazi opomnikov										
4.2.4.	Druge oblike kombiniranja načinov										

Legenda: IKT_GOS – Anketa o uporabi IKT v gospodinjstvih in posameznikih, AMP – Anketa o mnenju potrošnikov, TU-ČAP – Turistična potovanja domačega prebivalstva, SPK-AZK – Anketa o žrtvah kriminala, AIO – Anketa o izobraževanju odraslih, APEVG – Anketa o porabi energije in goriv v gospodinjstvih, KDZ – Kariera doktorjev znanosti, SILC – Raziskava o življenjskih pogojih, IKT_PO DJ – Uporaba IKT v podjetjih, PA-GRAD/M – Anketa o poslovnih tendencah v gradbeništvu

Tabela 4.3 : Podrobni podatki o raziskavah SURS

	Naslov raziskave	IKT_GOS		AMP	TU-ČAP	SPK-AZK		AIO		APEVG		KDZ		SILC		IKT_PODJ	PA-GRAD/M
	Leto	2008		2010	2010	2009		2007		2009		2010		Panel 2010		2009	2009
	Vzorčni okvir	Okvir popisnih okolišev in CRP		Telefonski imenik	Telefonski imenik, v prihodnosti CRP	CRP		CRP		/		SRDAP, POPIS 2002, RP		CRP		Poslovni register Slovenije	Poslovni register Slovenije
Najava in vabilo	Način najave	Poštno (/)		Poštno (/)	Poštno (/)	Poštno (/)		Poštno (/)		Poštno (/)		Poštno (/)		Poštno (/)		Poštno (/)	Poštno (/)
	Način vabila	Poštno (/)		Poštno (/)	Poštno (/)	Poštno (/)		Poštno (/)		Poštno (/)		Poštno (/)		Poštno (/)		Poštno (/)	Poštno (/)
	Strošek	/		/	/	/		/		/		/		/		/	/
Začetni način anketiranja	Način anketiranja	PAPI	CATI	CATI	CATI	CATI	CAPI	CATI	PAPI	CATI	PAPI	POŠTNA	CATI	CATI	CAPI	Poštno (ELEKTRONSKO 2005)	Poštno
	Povprečno trajanje ankete	13,8	9,8	5 minut 37 sekund	4 minute 48 sekund	19 minut		/		/		/		30 minut		20–30 minut	3 minute
	Način 1. opomnika	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/
	Način 2. opomnika	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/
	Začetna velikost (št. izbranih v vzorec)	1676	828	15000	15000	1084	916	7200		6000		1485		12500		2476	Letno povprečje: 513
	Kontaktiranih	1641	819	1486	14861	1075	874	/	/	/	/	/	/	12243		2458	Letno povprečje: 506
	Neustrezni	35	9	Letno povprečje: 14	139	9	42	/	/	/	/	/	/	257		18	Povprečje: 7
	Ustrezni	1641	434+385	1486	14861	1075	874	/	/	/	/	/	/	12243		2458	506
	Respondenti (končna velikost vzorca)	1092	386	719	8901	561	336	/	/	/	/	/	/	9028		2213	394
	Stopnja odgovora	66,50 %	47,10 %	Letna stopnja 48,4 %	59,90 %	52,20 %	38,40 %	/	/	/	/	/	/	73,74 %		95,5 % uteženo	89,6 % uteženo
	Strošek anketiranja	20.727 €	854 €	2383 delovnih ur	2562 delovnih ur	3340 delovnih ur		/	/	/	/	/	/	300.000 €		44.584 €	23.670 €

	Naslov raziskave	IKT_GOS		AMP	TU-ČAP	SPK-AZK		AIO		APEVG		KDZ		SILC		IKT_PODJ	PA-GRAD/M
	Leto	2008		2010	2010	2009		2007		2009		2010		Panel 2010		2009	2009
Menjava načina anketiranja za neuspešno anketirane z začetnim načinom	Način anketiranja		PAPI	NI DODATNEGA ANKETIRANJA	NI DODATNEGA ANKETIRANJA		CAPI		PAPI		/		/		CAPI	/	CATI v primeru manjkajočih odgovorov
	Povprečno trajanje ankete		14 minut	NE	NE		19 minut		/		/		/	30 minut		/	/
	Način 1. opomnika		/	NE	NE		/		/		/	/	/		/	/	/
	Način 2. opomnika		/	NE	NE		/		/		/	/	/		/	/	/
	Začetna velikost (št. izbranih v vzorec)		385	NE	NE		357		/		/	/	/		/	/	/
	Neustrezni		3	NE	NE		8		/		/	/	/		/	/	/
	Ustrezni		382	NE	NE		349		/		/	/	/		/	/	/
	Respondenti (končna velikost vzorca)		223	NE	NE		117		/		/	/	/		/	/	/
	Stopnja odgovora		58,40 %	NE	NE		33,52 %		/		/	/	/		/	/	/
	Strošek anketiranja		20.727 € (skupaj s prvim)	NE	NE		/		/		/	/	/		/	/	/

Legenda: / – ni podatka, IKT_GOS – Anketa o uporabi IKT v gospodinjstvih in posameznikih, AMP – Anketa o mnenju potrošnikov, TU-ČAP – Turistična potovanja domačega prebivalstva, SPK-AZK – Anketa o žrtvah kriminala, AIO – Anketa o izobraževanju odraslih, APEVG – Anketa o porabi energije in goriv v gospodinjstvih, KDZ – Kariera doktorjev znanosti, SILC – Raziskava o življenjskih pogojih, IKT_PODJ – Uporaba IKT v podjetjih, PA-GRAD/M – Anketa o poslovnih tendencah v gradbeništvu

Iz zgornjih tabel opazimo, da se uporaba kombiniranih načinov v anketah SURS nanaša predvsem na uporabo različnega načina za najavo ankete glede na zbiranje podatkov ter na kombinacijo načinov v fazi zbiranja podatkov in fazi ponovnega anketiranja neodgovorov. V SURS-ovih raziskavah se pojavita dva tipa kombiniranja načinov. Prvi se uporablja v anketah gospodinjstev, drugi tip pa je značilen za ankete podjetij. SURS uporablja različne načine kombiniranja v uradnih anketah zaradi administrativnih razlogov ter za dvigovanje stopnje odgovora, ne pa tudi zaradi optimizacije stroškov. Za vzorčni okvir se najpogosteje uporablja telefonski imenik in CRP. Ker smo priča trendu upadanja gospodinjstev s fiksnim telefonom in padcu objavljanja telefonskih števil v registrih, bi bilo potrebno razmisliti o drugih vzorčnih okvirih namesto telefonskega imenika. SURS vse svoje ankete najavlja preko pisemske pošte. Najpogosteje uporabljeni metodi zbiranja podatkov sta telefonsko (CATI) in terensko (CAPI) anketiranje. Ob naraščajoči penetraciji interneta in naraščajočih stroških dela bi bilo potrebno več pozornosti nameniti spletnim anketam in uporabi več različnih kombiniranih načinov z uporabo spleta.

5 Empirični del

V empiričnem delu želimo ugotoviti, kako se spreminja odnos med stroški in napakami med telefonskim in terenskim zbiranjem podatkov.

5.1 Opis raziskave

Anketa o uporabi informacijsko-komunikacijske tehnologije v gospodinjstvih in pri posameznikih se na Statističnem Uradu Republike Slovenije izvaja od leta 2004. SURS z raziskovanjem ugotavlja, kako razširjena je uporaba računalnikov in drugih informacijsko-komunikacijskih naprav, kolikšen delež slovenske populacije dostopa do interneta in za katere namene ga uporablja. Za omenjeno raziskavo smo se odločili zaradi hitre širitve interneta in vedno večje uporabe IKT. V empiričnem delu diplomske naloge bomo operirali s podatki ankete o uporabi informacijsko-komunikacijske tehnologije v gospodinjstvih in pri posameznikih 2008 (IKT_GOS_2008). Podatki so bili leta 2008 pridobljeni z uporabo telefonskega in terenskega načina zbiranja podatkov.

5.2 Opis vzorca

Osnova za vzorčni okvir za raziskovanje sta bila okvir popisnih okolišev in Centralni register prebivalstva (CRP). Vzorec je bil stratificiran dvostopenjski (s ponavljanjem), eksplicitno stratificiran po velikosti in tipu naselja, implicitno pa po statističnih regijah, hkrati je proporcionalen glede na število izbranih oseb v vsakem stratumu. Na prvi stopnji so izbrali 313 vzorčnih enot (skupin popisnih okolišev), nato pa v vsaki izbrani vzorčni enoti po 8 oseb, starih med 10 in 74 let. Vzorec je tako obsegal 2504 oseb.

Izbrane vzorčne enote imenujemo segmenti in so oštevilčeni od 1 do 313, izbrane osebe pa so oštevilčene od 1 do 2504. Številka izbrane osebe je enolična identifikacija v datoteki.

Verjetnost izbora osebe se izračuna po formuli

$$\frac{N_a}{\sum_{S10} N_a} \cdot \frac{8}{N_a},$$

kjer je N_a število oseb, starih med 10 in 74 let v vzorčni enoti. Izbrali so tiste osebe, ki so na dan začetka anketiranja že dopolnile 10 let in ob koncu obdobja anketiranja še niso presegle starosti 74 let (Belak in Zdešar 2008, 2).

5.3 Prilagoditev vzorca

Iz baze podatkov izločimo vse osebe, ki so starejše od 50 let. V prilagojenem vzorcu je 1408 oseb starih od 10 do 50 let. V podvzorec za telefonsko zbiranje podatkov je bilo izbranih 492, v podvzorec za terensko anketiranje pa 916 oseb. Osebe, ki niso bile dostopne ali niso hotele sodelovati v telefonski anketi, so anketarji anketirali na terenu. V začetni vzorec telefonskega načina anketiranja je bilo zbranih 492 enot, od tega je bilo 5 enot neustreznih. Stopnja odgovora je bila 40,9 %, kar pomeni, da je v telefonski anketi preko fiksnega telefona sodelovalo 192 ljudi, preko mobilnega telefona pa je sodelovalo 9 ljudi, skupaj je v telefonskem anketiranju sodeloval 201 respondent. V začetni vzorec terenskega anketiranja je bilo izbranih 916 oseb, od tega je bilo 18 enot neustreznih. Stopnja odgovora je bila 76,90 %, kar pomeni, da je v terenski anketi sodelovalo 691 ljudi.

5.4 Izbira spremenljivk

Najprej je potrebno izbrati ključne spremenljivke raziskave, jih po potrebi rekodirati v dihotomne spremenljivke ter jim določiti statistike za posamezen način zbiranja podatkov.

Tabela 5.1: Izbira in rekodacija spremenljivk

ID	Opis	Rekodacija
A2	Dostop do interneta v gospodinjstvu	0 = Nima dostopa 1 = Ima dostop
C1	Uporaba interneta	0 = Ne uporablja 1 = Uporablja
C2	Pogostost uporabe interneta	0 = Ne uporablja ali redko uporablja 1 = Pogosto
D8	Uporaba mobilnega telefona	0 = Ne uporablja 1 = Uporablja
B1	Uporaba računalnika	0 = Ne uporablja 1 = Uporablja
B2	Pogostost uporabe računalnika v zadnjih 3 mesecih	0 = Ne uporablja ali redko uporablja 1 = Pogosto
D6_b	Potencialni razlogi za nakup ... pravica do izmenjave zakonsko zaščitene vsebine	0 = Ne 1 = Da
D6_e	Potencialni razlogi za nakup ... plačljive vsebine kakovostnejše od brezplačnih	0 = Ne 1 = Da
D7_c	Nadomestitev kupovanja ali izposojanja DVD-jev s prenašanjem filmov in video posnetkov z interneta	0 = Nisem nadomestil 1 = Sem nadomestil
A1	Število naprav IKT v gospodinjstvu	COMPUTE
G4	Mesečni neto dohodek gospodinjstva	/
F1	Stopnja izobrazbe	0 = Srednja šola ali manj 1 = Višja šola ali več
DS	Dejanska starost	/

5.5 Metoda, analiza in interpretacija

Za vsako spremenljivko bomo izračunali pravo vrednost oziroma aritmetično sredino (μ) ali vrednost, ki jo dobimo, če upoštevamo telefonsko in terensko anketiranje ter standardni odklon μ . Kot pravo vrednost obravnavamo vrednost, ki jo dobimo iz načina, za katerega menimo, da daje najbolj natančno oceno oziroma najbližjo oceno dejanski populacijski vrednosti. Iz baze je potrebno izločiti vse tiste enote, ki so bile anketirane terensko, in nato vse tiste, ki so bile anketirane telefonsko; tako bomo lahko izračunali aritmetično sredino oziroma vrednost telefonskega in terenskega anketiranja ter standardni odklon. Vsem spremenljivkam je potrebno izračunati varianco. Za dihotomne spremenljivke (tiste z vrednostjo 0 in 1) bomo uporabili formulo variance ocene deležev, in sicer za izračun variance prave vrednosti formulo:

$$VAR(\mu) = \frac{\mu * (1 - \mu)}{(n_{vsi} - 1)},$$

za izračun variance dobljene vrednosti s posameznim načinom pa:

$$VAR(\bar{X}) = \frac{\bar{X} * (1 - \bar{X})}{(n_{način} - 1)}.$$

Za številske spremenljivke (dohodek, število IKT-naprav in starost) pa bomo uporabili formulo variance ocene povprečji. Za izračun prave vrednosti oziroma vrednosti, dobljene s telefonskim in terenskim anketiranjem, bomo uporabili formulo:

$$VAR(\mu) = \frac{s_{\bar{\mu}}^2}{n_{vsi}},$$

za izračun variance dobljene vrednosti posameznega načina pa:

$$VAR(\bar{X}) = \frac{s_{\bar{X}}^2}{n_{način}}.$$

Vrednost pristranosti se izračuna kot razlika med pravo in dobljeno vrednostjo:

$$Bias(\bar{X}) = \bar{X} - \mu,$$

Vrednost srednje kvadratne napake se izračuna kot vsota kvadrata pristranosti in variance:

$$MSE = Bias(\bar{X})^2 + Var(\bar{X})$$

Biemer (2010, 840) v svojem članku o skupni anketni napaki izpostavi natančnejšo formulo za izračun vrednosti srednje kvadratne napake, pri kateri je prava vrednost dobljena na osnovi druge ankete in s katero bomo operirali v nadaljnji analizi:

$$MSE(\bar{X}) = Bias(\bar{X})^2 - Var(\mu) + 2 \times \sqrt{Var(\bar{X}) \times Var(\mu)}.$$

MSE je lahko različen za različne spremenljivke, zato je potrebno opazovati več spremenljivk hkrati. Za optimizacijo napak in stroškov se v raziskovanju o stratifikaciji in vzorčenju v skupinah uporablja postopek minimizacije produkta vzorčne variance in stroškov, kjer je optimalnejša izbira tista z nižjim produktom (Cochran 1978). Sodobnejši pristopi so vzorčno varianco nadomestili s srednjo kvadratno napako (Vehovar in drugi 2010).

Pri določanju razmerja med stroški in napakami se lahko odločimo za tri možnosti (Vehovar in drugi, 2010):

1. načrt, ki zahteva fiksni MSE opazovane spremenljivke in ki ima najnižje stroške; ta pristop se uporablja pri merjenjih brezposelnosti,
2. načrt z najnižjim MSE opazovane spremenljivke ob fiksnih stroških; ta pristop se uporablja pri raziskavah z omejenim proračunom,
3. najmanjši produkt stroškov in MSE za opazovano spremenljivko; ta pristop je najbolj fleksibilen.

V analizi bomo uporabili tretji pristop. Da bomo lažje ocenili, kateri način je boljši za posamezno spremenljivko, je potrebno narediti še simulacijo z različnimi velikostmi vzorca.

Izbrali smo sedem različnih velikosti vzorca (100, 200, 500, 1000, 2000, 5000 in 10000), nato bomo ob predpostavkah, da se stopnja pristranskosti ne spreminja in da stopnja odgovora pri posameznem načinu ostane ista kot je v osnovi (41 % CATI in 77 % CAPI), za vsak par spremenljivk izračunali produkt MSE in stroškov. Kot zanimivost pa bomo v zadnjem delu empiričnega dela naloge naredil isto simulacijo v primeru nagrad respondentom. Predpostavljali bomo, da se stopnja odgovora pri terenskem načinu poveča za 5 odstotnih točk, pri telefonskem pa za 20 odstotnih točk, če v primeru sodelovanja po telefonu ponudimo nagrado 10 €, v primeru sodelovanja na terenu pa 5 €.

Tabela 5.2: MSE in stroški, telefonsko anketiranje

Spremenljivke	Določitev statistik					Sodelovanje						MSE in stroški			
	μ	Std. dev μ	Varianca μ	$\bar{X}$	Std. dev $\bar{X}$	Začetni n	Ustrezni	Odgovori (%)	Odgovori (n)	Priistranost (Bias)	Varianca $\bar{X}$	MSE	Stroški	MSE x stroški	Strošek na respondenta
Dostop do interneta	0,82	0,381	0,00017	0,82	0,38	492	487	40,90 %	201	0,000	0,00074	0,00054	1.190,52 €	0,64	5,92 €
Uporaba interneta	0,77	0,42	0,0002	0,776	0,42					0,006	0,0009	0,0007		0,83	
Pogostost uporabe interneta	0,755	0,43	0,0002	0,782	0,41					0,027	0,0008	0,0013		1,55	
Uporaba mobilnega telefona	0,97	0,165	0,000033	0,98	0,16					0,01	0,0001	0,00018		0,21	
Uporaba računalnika	0,801	0,4	0,00018	0,814	0,39					0,013	0,0008	0,00074		0,88	
Pogostost uporabe računalnika	0,775	0,41	0,0002	0,808	0,4					0,03	0,0008	0,0015		1,78	
Pravica do izmenjave zakonsko zaščitene vsebine	0,436	0,49	0,00028	0,522	0,501					0,086	0,0013	0,0083		9,88	
Plačljive vsebine kakovostnejše od brezplačnih	0,51	0,50	0,00029	0,594	0,492					0,084	0,0012	0,0079		9,40	
Nadomestitev kupovanja s prenašanjem vsebin iz interneta	0,364	0,48	0,00026	0,42	0,494					0,056	0,0012	0,004		4,76	
Izobrazba	0,223	0,42	0,00019	0,227	0,42					0,004	0,001	0,00068		0,81	

Tabela 5.3: MSE in stroški za številske spremenljivke, telefonsko anketiranje

Spremenljivke	Določitev statistik					Sodelovanje						MSE in stroški			
	μ	Std. dev μ	Varianca μ	$\bar{X}$	Std. dev $\bar{X}$	Začetni n	Ustrezni	Odgovori (%)	Odgovori (n)	Priistranost (Bias)	Varianca $\bar{X}$	MSE	Stroški	MSE x stroški	Strošek na respondenta
Število IKT-naprav	4,23	0,97	0,0011	4,4	0,93	492	487	40,90 %	201	0,17	0,0043	0,032	1.190,52 €	38,1	5,92 €
Dohodek	1627,350	924,6	958,34	1703,7	1056,28					76,33	5550,88	9480,8		11287070	
Starost	34,73	9,46	0,1	34,65	9,6					-0,08	0,46	0,34		404,77	

Tabela 5.4: MSE in stroški, terensko anketiranje

Spremenljivke	Določitev statistik					Sodelovanje						MSE in stroški			
	μ	Std. dev μ	Varianca μ	$\bar{X}$	Std. dev $\bar{X}$	Začetni n	Ustrezni	Odgovori (%)	Odgovori (n)	Pristranost (Bias)	Varianca $\bar{X}$	MSE	Stroški	MSE x stroški	Strošek na respondenta
Dostop do interneta	0,82	0,381	0,00017	0,822	0,38	916	898	76,90 %	691	0,002	0,00021	0,0002	12.091,45 €	2,41829	17,50 €
Uporaba interneta	0,77	0,42	0,0002	0,75	0,428					-0,02	0,00027	0,0007		8,464015	
Pogostost uporabe interneta	0,755	0,43	0,0002	0,752	0,432					-0,003	0,00027	0,0003		3,627435	
Uporaba mobilnega telefona	0,97	0,165	0,000033	0,97	0,171					0,000	0,00004	0,00004		0,483658	
Uporaba računalnika	0,801	0,4	0,00018	0,79	0,407					-0,011	0,00024	0,0004		4,83658	
Pogostost uporabe računalnika	0,775	0,41	0,0002	0,76	0,382					-0,015	0,00026	0,00048		5,803896	
Pravica do izmenjave zakonsko zaščitene vsebine	0,436	0,49	0,00028	0,383	0,487					-0,053	0,00034	0,0032		38,69264	
Plačljive vsebine kakovostnejše od brezplačnih	0,51	0,50	0,00029	0,452	0,498					-0,058	0,00036	0,0037		44,73837	
Nadomestitev kupovanja s prenašanjem vsebin iz interneta	0,364	0,48	0,00026	0,345	0,475					-0,019	0,00033	0,0007		8,464015	
Izobrazba	0,223	0,42	0,00019	0,22	0,41					-0,003	0,00025	0,00025		3,022863	

Tabela 5.5: MSE in stroški za številske spremenljivke, terensko anketiranje

Spremenljivke	Določitev statistik					Sodelovanje						MSE in stroški			
	μ	Std. dev μ	Varianca μ	$\bar{X}$	Std. dev $\bar{X}$	Začetni n	Ustrezni	Odgovori (%)	Odgovori (n)	Pristranost (Bias)	Varianca $\bar{X}$	MSE	Stroški	MSE x stroški	Strošek na respondenta
Število IKT-naprav	4,23	0,97	0,0011	4,19	0,99	916	898	76,90 %	691	0,04	0,0014	0,003	12.091,45 €	36,27435	17,50 €
Dohodek	1627,350	924,6	958,34	1536,4	835,36					-90,96	1009,88	9282,9		112243721,2	
Starost	34,73	9,46	0,1	34,76	9,54					0,03	1,75	0,737		8911,39865	

Dobljene rezultate najprej interpretiramo vsebinsko, glede na razlike med pravimi in dobljenimi vrednostmi. Ljudje, ki so bili anketirani preko telefona, imajo pričakovano višji dohodek od tistih, ki ga nimajo, poleg tega imajo ljudje, ki so bili anketirani preko telefona, več IKT-naprav, bolj pogosto uporabljajo računalnik, več kupujejo IKT zaradi pravice do izmenjave zakonsko zaščitene vsebin, bolj nadomeščajo nakupe s prenašanjem internetnih vsebin, hkrati pa se bolj strinjajo s trditvijo, da so plačljive stvari kakovostnejše.

Stroški telefonskega anketiranja so 1.190,52 €, kar pomeni, 5,92 € na respondenta. V tabeli 5.2 vidimo, da ima največjo vrednost MSE pri dihotomnih spremenljivkah spremenljivka »pravica do izmenjave zakonsko zaščitene vsebin«, in sicer 0,0083. Sledi ji spremenljivka »plačljive vsebine so kakovostnejše od brezplačnih«, ki ima vrednost 0,0079. Spremenljivka »nadomestitev kupovanja s prenašanjem vsebin iz interneta« ima vrednost 0,004, vse ostale spremenljivke pa imajo nižje vrednosti. Produkt MSE in stroškov je pričakovano najvišji pri spremenljivkah, ki imajo višji MSE. Pri spremenljivki z najvišjim MSE (pravica do izmenjave zakonsko zaščitene vsebin) je produkt 9,88. Najnižji produkt v telefonskem anketiranju ima spremenljivka »uporaba mobilnega telefona«, in sicer 0,21.

MSE za številske spremenljivke zaradi različnih merskih lestvic ne moremo primerjati med seboj. Produkt pri spremenljivki »število IKT-naprav« je 38,1, pri spremenljivki »dohodek« 11287070, pri starosti pa je produkt 404,77.

Iz Tabele 5.4 vidimo, da so bili stroški terenskega anketiranja 12.091,45 € oziroma 17,50 € na respondenta. Največjo vrednost MSE pri dihotomnih spremenljivkah ima spremenljivka »plačljive vsebine so kakovostnejše od brezplačnih«, in sicer 0,0037, medtem ko ima najnižjo vrednost MSE »uporaba mobilnega telefona« 0,00004. Najvišji produkt stroškov in MSE je pri spremenljivki z najvišjo vrednostjo MSE, 44,74. Sledi ji spremenljivka »pravica do izmenjave zakonsko zaščitene vsebin« s produktom 38,7, medtem ko ima spremenljivka z najnižjim MSE vrednost 0,48. Spremenljivka »število IKT-naprav« ima produkt 36,27, »dohodek« 112243721,1 in starost 8911,4.

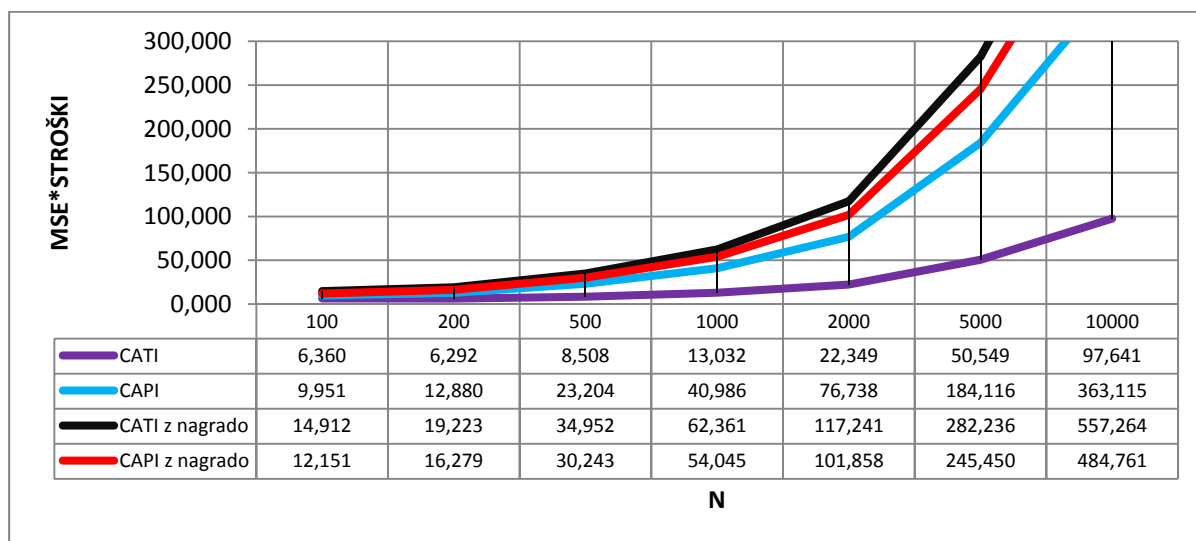
Na podlagi vrednosti produktov MSE in stroškov lahko rečemo, da je za vse izbrane spremenljivke, razen za spremenljivko »število IKT-naprav«, optimalnejši telefonski način zbiranja podatkov, ker je vrednost le-tega nižja kot pri terenskem načinu zbiranja podatkov.

Največja razlika med načinoma je pri spremenljivki »uporaba interneta«, kjer je produkt pri terenskem desetkrat višji (8,46) kot produkt pri telefonskem načinu (0,83), in pri spremenljivki »plačljive vsebine kakovostnejše od brezplačnih«, kjer je vrednost pri telefonskem načinu 9,4, pri terenskem načinu pa 44,74. Omenjeni spremenljivki sodita v kategorijo vprašalnika, kjer so navedeni potencialni razlogi za nakup, torej lahko zaključimo, da bi bilo glede na te rezultate optimalnejše postavljati omenjeno vprašanje preko telefona kot pa na terenu.

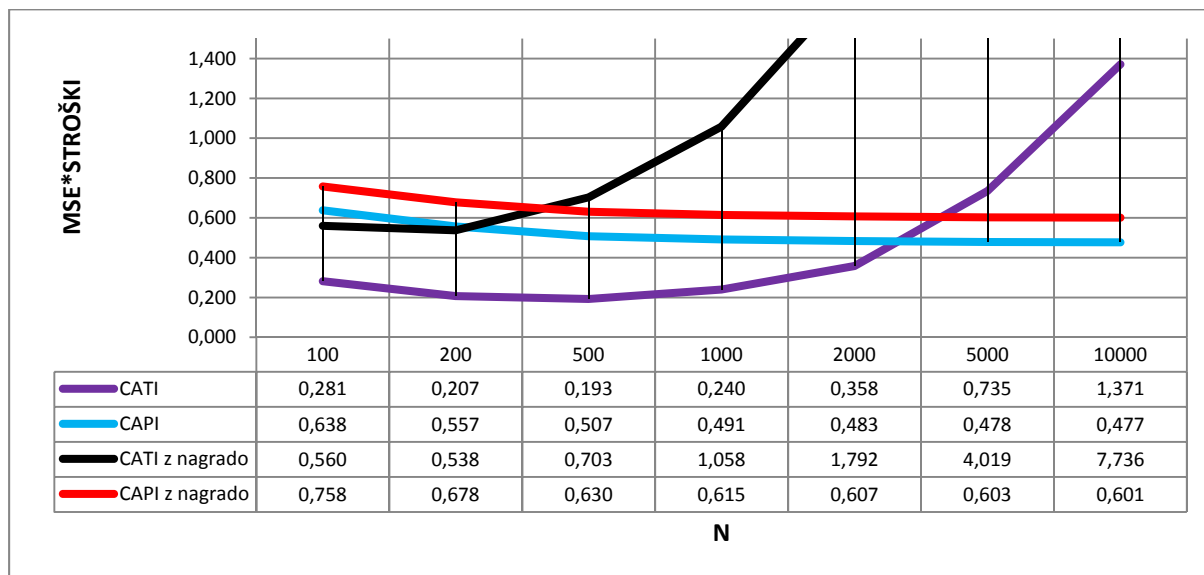
Iz rezultatov je razvidno, da je v našem primeru optimalnejši telefonski način zbiranja podatkov, vendar če pogledamo dovolj široko in poleg produkta stroškov in MSE vzamemo v obzir tudi stopnje odgovorov, vidimo, da je bila stopnja odgovora pri terenskem načinu za faktor 1,88 večja od telefonske ankete. Nadalje je razvidno, da je bilo na terenu dejansko anketiranih 691 respondentov, medtem ko je bil telefonsko anketiran 201 respondent. To pomeni, da je bilo v osnovi zaradi večjega vzorca terensko anketiranje dražje od telefonskega.

Da bomo lahko dejansko trdili o boljšem ali slabšem načinu zbiranja podatkov, je potrebno narediti še simulacijo med načinoma glede na velikost vzorca pri posameznem vprašanju. Omenjeni postopek bomo ponovili še v primeru denarnih nagrad za sodelovanje pri anketi. Terenskim respondentom damo 5 €, telefonskim pa 10 € nagrade. Predpostavljamo, da se stopnja odgovora pri telefonski anketi poveča za 20 %, pri terenski pa za 5 %.

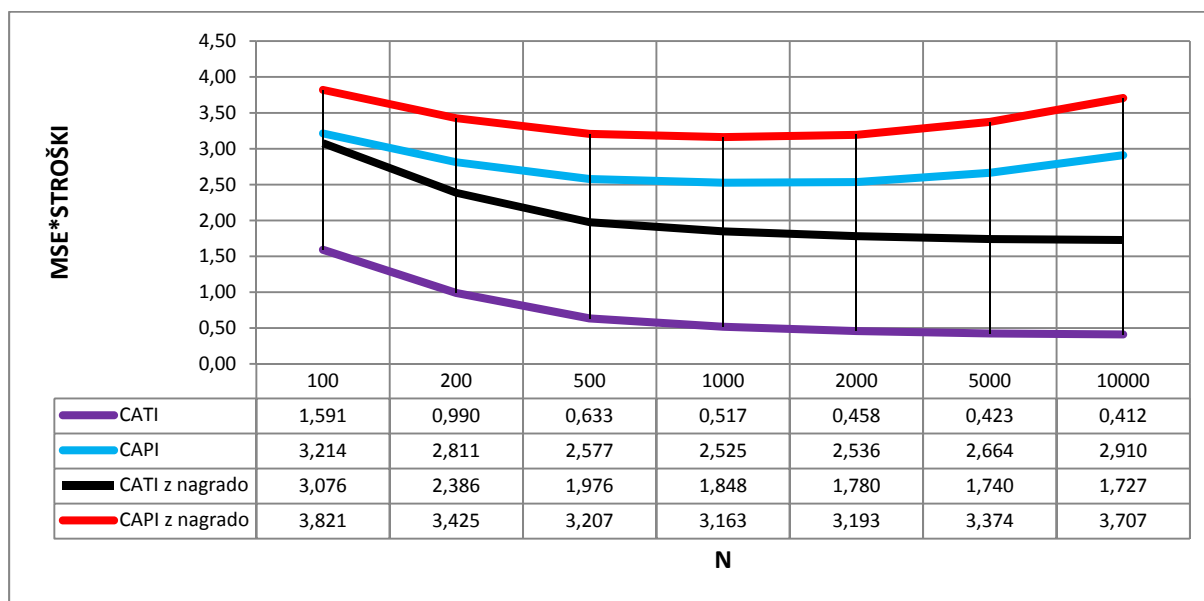
Slika 5.1: Simulacija spremenljivke »pravica do izmenjave zakonsko zaščitene vsebine« brez nagrade in z nagrado



Slika 5.2: Simulacija spremenljivke »uporaba mobilnega telefona« brez nagrade in z nagrado



Slika 5.3: Simulacija spremenljivke »dostop do interneta« brez nagrade in z nagrado

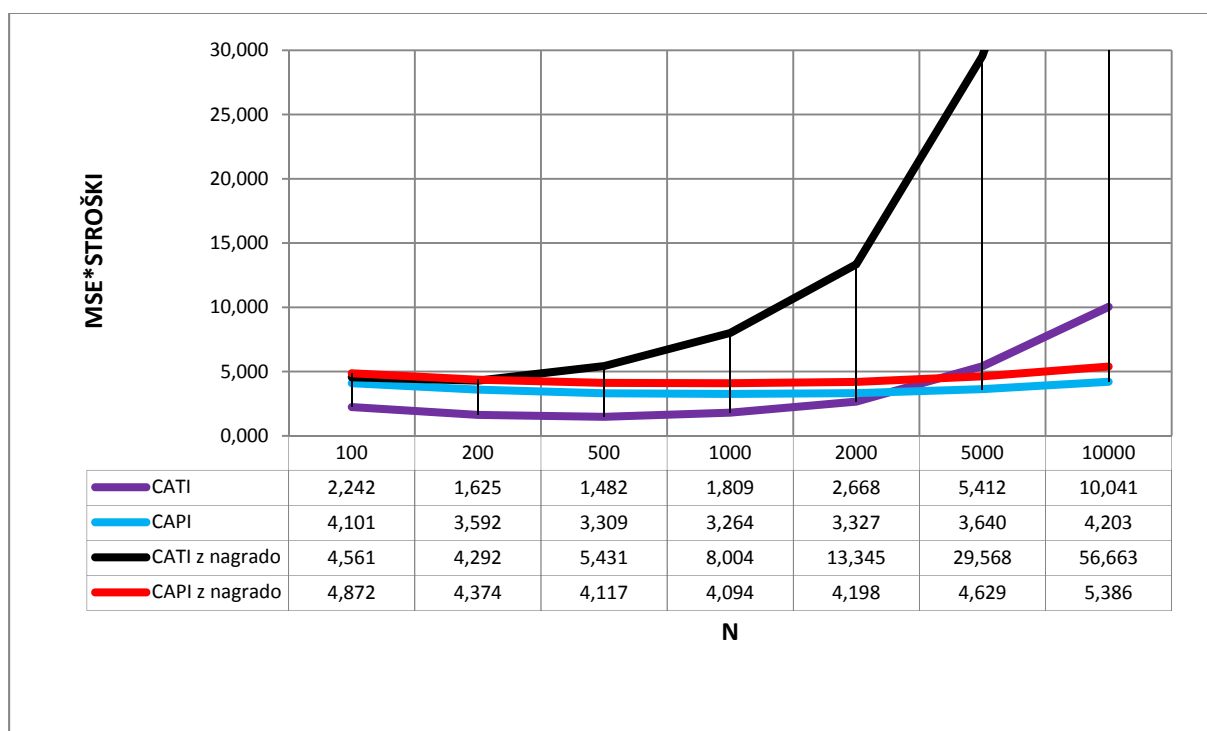


Po opravljenih simulacijah (glej Priloga B) opazimo, da so krivulje spremenljivk različne. V splošnem lahko opazovane spremenljivke razdelimo v tri skupine:

1. v skupino spremenljivk, kjer se s povečevanjem velikosti vzorca v vseh načinih povečuje tudi produkt (Slika 5.1): »uporaba interneta«, »pravica do izmenjave zakonsko zaščitene vsebine«, »plačljive vsebine so kakovostnejše od brezplačnih«, »nadomestitev kupovanja s prenašanjem vsebin iz interneta« in »dohodek«,

2. v skupino spremenljivk, kje se krivulje načinov sekajo pri različnih velikostih vzorca (Slika 5.2): »pogostost uporabe interneta«, »uporaba mobilnega telefona«, »pogostost uporabe računalnika«, »število IKT-naprav« in »starost«,
3. v skupino spremenljivk, kjer je ne glede na velikost vzorca najoptimalnejši en način zbiranja podatkov, kjer se krivulje načinov ne sekajo, produkti posameznega načina se nižajo ali povečujejo (Slika 5.3): »dostop do interneta«, »uporaba interneta«, »uporaba računalnika« in »izobrazba«.

Slika 5.4: Simulacija spremenljivke »pogostost uporabe interneta« brez nagrade in z nagrado



Podrobneje bomo opisali simulacijo spremenljivke »pogostost uporabe interneta«, ker je simulacija zelo ilustrativna in vsebuje vse elemente zgoraj opisanih skupin spremenljivk (Slika 5.1, Slika 5.2 in Slika 5.3). Iz Slike 5.4 opazimo, da je do velikosti vzorca 2000 enot najoptimalnejši CATI-način zbiranja podatkov, od velikosti vzorca 5000 enot pa je optimalnejši CATI-način tako brez kot z nagrado. Iz grafa opazimo, da se krivulja produkta CATI-načina do velikosti vzorca 500 enot spušča. Na tem odseku (100–500 enot) ima CATI največji produkt pri velikosti 100 enot, in sicer 2,242. Predpostavljamo, da je vrednost pristranskosti pri vseh velikosti vzorca enaka, medtem ko se vrednost variance skozi velikosti vzorca spreminja, zato je višji začetni produkt posledica visokih fiksnih stroškov. Varianca pri

velikosti vzorca 100 enot je 0,0042, medtem ko je pri velikosti 500 enot 0,00837. Tako lahko trdimo, da se produkt kljub višjim stroškom pri večjih velikosti manjša zaradi variance. Pri velikosti 1000 enot začne krivulja hitro naraščati. Stroški se z večanjem vzorca sicer povečujejo, varianca na drugi strani pa se manjša, vendar ne tako hitro kot pri nizkih vzorcih, zato začne krivulja naraščati. Prav tako kot pri krivulji CATI je začetni produkt večji tudi pri CAPI-načinu. Pri velikosti 100 enot je produkt 4,101 zaradi visokih fiksnih stroškov, medtem ko je varianca 0,00245. Produkt z večanjem vzorca pada vse do velikosti 2000 enot. Pri večjih vzorcih se začne produkt sicer povečevati, vendar počasneje kot pri CATI-načinu pri enako velikih vzorcih, zato pride pri velikosti 5000 enot do zamenjave optimalnejšega načina. Produkt pri CATI je 5,412, pri CAPI pa 3,640. Pri načinih z nagrado je situacija nekoliko drugačna, saj CAPI-način z nagrado postane optimalnejši že pri vzorcu 500 enot. Pri manjših vzorcih (100–500) sta si oba načina z nagrado približno enakovredna. Pri CATI-načinu z nagrado je varianca pri velikosti 100 enot 0,002841, kar pomeni, da je varianca nižja od tiste pri načinu brez nagrade (0,0042), zato lahko rečemo, da je visok začetni produkt MSE in stroškov posledica visokih začetnih stroškov. Varianca CAPI-načina z nagrado je pri velikosti 100 enot 0,002302, pri načinu brez nagrade pa je 0,002457. Variance načinov so si skozi vse vzorce zelo podobne, zato sta si krivulji CAPI-načina z in brez nagrade na sliki zelo blizu ne glede na velikost vzorca. Če primerjamo variance načinov z nagrado, opazimo, da se med sabo manj razlikujejo kot variance načinov brez nagrade, vendar je iz slike vseeno razvidno, da je razlika med načinoma veliko večja kot pri načinoma brez nagrade. Ta pojav lahko pojasnimo kot posledico nagrajevanja (terenskim respondentom smo dali nagrado 5 €, telefonskim pa 10 €), zato so razlike med načinoma toliko večje.

5.6 Omejitve

Iz rezultatov empiričnega dela vidimo, da se optimalni načini z večanjem vzorca pri spremenljivkah razlikujejo. Razlike so vzrok med spremenljivkami neenakomerno padajoče variance in visokih fiksnih stroškov posameznega načina. Glavna omejitev empiričnega dela je predpostavka, da se z večanjem vzorca cena posamezne spremenljivke ne spreminja; spreminja se le vzorčna varianca, ki se z večanjem vzorca manjša. Vzorčna varianca se pri velikih vzorcih ne zmanjšuje tako hitro kot pri majhnih vzorcih, medtem ko stroški z večanjem vzorca hitro naraščajo. V primeru, da bi se spreminjala tudi ocena (s povečevanjem vzorca bi bila natančnejša), bi se zmanjšal tudi produkt. Druga omejitev empiričnega dela so neuteženi podatki po demografskih spremenljivkah. V primeru uteževanja bi bili rezultati seveda lahko drugačni od dobljenih. Poleg uteževanja bi bilo priporočljivo ponoviti postopek ločeno za glavno fazo zbiranja in za nadomestno fazo zbiranja podatkov. V IKT_GOS 2008 so v prvem valu anketirali z načinom CATI in CAPI, nerespondente v CATI podvzorcu pa so v primeru neodgovorov anketirali še na terenu. Segment respondentov, ki bi ga bilo priporočljivo opazovati ločeno, so torej nerespondenti na prvi način. V empiričnem delu smo torej vse tiste, ki so bili anketirani na terenu, uvrstili v svoj segment.

6 Zaključek

V diplomski nalogi smo predstavili različne načine zbiranja podatkov, njihove prednosti in slabosti, stroške ter napake, ki se pojavljajo v posameznem načinu. V glavnem delu diplomske naloge smo predstavili kombiniranje načinov anketiranja ter opisali najpogosteje uporabljene kombinacijske načine razširjene klasifikacije po de Leeuw (2005). Številne možnosti kombiniranja načinov v vseh fazah raziskovalnega procesa ponujajo raziskovalcem veliko možnosti vplivanja na kvaliteto podatkov, hkrati pa imajo raziskovalci glede na omejena finančna sredstva več nadzora nad stroški raziskave. Kombiniranje načinov pa nima samo pozitivnih lastnosti, kajti nepremišljeno kombiniranje lahko negativno vpliva na stroške in na kvaliteto podatkov.

S pregledom raziskav SURS smo ugotovili, da le-ta uporablja kombiniranje načinov predvsem v fazi kontaktiranja v primerjavi s fazo zbiranja podatkov (način najave ankete je drugačen od načina zbiranja podatkov). Poleg tega kombinirajo tudi v glavni fazi zbiranja podatkov v primerjavi s fazo opomnikov; v slednji uporabijo dražji način iz predhodne faze.

V empiričnem delu smo se osredotočili na raziskavo SURS o IKT v gospodinjstvih, ki uporablja CATI in CAPI, ter skušali odgovoriti na raziskovalno vprašanje: *Kako se spreminja odnos med stroški in napakami med načini anketiranja posamezne spremenljivke z večanjem vzorca?* Ugotovili smo, da se v številnih primerih produkt napak in stroškov z večanjem vzorca večja. Produkt se v veliki meri povečuje zaradi stroškov, medtem ko ocena tudi z večanjem vzorca ostaja nespremenjena. Pri skoraj vseh spremenljivkah (razen pri spremenljivki »število IKT-naprav«) je pri majhnih vzorcih zato optimalnejši cenejši CATI način zbiranja podatkov. Z večanjem vzorca se razlike produktov pri nekaterih spremenljivkah zelo povečajo; v nekaterih primerih se začnejo razlike manjšati, v nekaterih primerih pa se krivulje posameznega načina tudi sekajo. Če se odločimo za kombinacijo načinov, moramo opazovati razmerje stroškov med posameznim načinom. V primeru načina brez nagrade (CATI in CAPI) se razmerje z večanjem vzorca sicer povečuje, vendar pri velikih vzorcih razmerje ne narašča tako hitro kot pri majhnih vzorcih. V primeru spremenljivke »nadomestitev kupovanja s prenašanjem vsebin preko interneta« in »starost« pa razmerje najprej narašča, nato pa se z večanjem vzorca manjša. V primeru nagrade smo priča istemu pojavu tudi pri spremenljivki »uporaba računalnika«, kjer najprej razmerje raste,

nato pa se začne pri velikih vzorcih ponovno manjšati. Posebej pri spremenljivki »pravica do izmenjave zakonsko zaščitene vsebine« se z večanjem vzorca razmerje hitro manjša. S simulacijo načinov z nagrado smo dosegli, da postaneta načina v nekaterih primerih primerljiva bistveno hitreje (pri manjših velikostih). Z večanjem vzorca se torej spreminja tudi odnos produkta stroškov in napak posameznega načina zbiranja podatkov. Pričakovali bi, da se bo z večanjem vzorca pri vseh spremenljivkah večal produkt, vendar temu ni tako zaradi zmanjševanja variance. V nekaterih primerih se produkt manjša, v nekaterih primerih pa se produkt enega načina z večanjem povečuje, produkt drugega pa manjša.

Za večino spremenljivk je najoptimalnejši telefonski način. Če se odločimo za kombiniranje načinov, pa je najoptimalnejša kombinacija telefonskega in terenskega načina. Pri nekaterih spremenljivkah (Slika B.1, Slika B.2, Slika B.6 in Slika B.8) postane telefonski način z nagrado optimalnejši od terenskega načina brez nagrade. Optimalno izbiro načinov za celotno anketo bi lažje določili, če bi imeli standardizirani MSE. Optimalno kombinacijo načinov dosežemo tako, da opazujemo krivulje posameznega načina. Velikost vzorca posameznega načina pri kombiniranju bi v našem primeru določili tako, da bi izbrali posamezen način na podlagi velikosti produkta stroškov in napak.

Z empiričnim delom smo pokazali, kako je potrebno gledati na razmerje MSE in stroškov, da ugotovimo, kateri način je optimalnejši. Statistični urad Republike Slovenije bi lahko začel preučevati optimizacijo stroškov in napak na podoben način in s tem tudi bolj uveljavljati kombiniranje načinov v svoje ankete. Z vse pogostejšo uporabo spletnih anket bi bilo potrebno na uradu razmisliti tudi v smeri uporabe in preučevanja spletnih anket. Za nadaljnje raziskovanje te problematike bi bilo potrebno za več ključnih spremenljivk standardizirati MSE, šele potem bi lahko iz rezultatov videli, katera metoda ali kombinacija metod je optimalna. Poleg tega bi bilo za primerjave potrebno podatke vsakič ločeno utežiti po demografskih spremenljivkah.

Eksperimentalno bi lahko istočasno izvedli neodvisno anketo, ki bi bila identična anketi SURS in na enakovrednem vzorcu enot kot pri uradni anketi. Eksperimentalni skupini bi tako poslali pošto vabilo za sodelovanje v spletni anketi, nerespondentom predhodne faze bi poslali poštni opomnik z možnostjo telefonskega anektiranja (call-in), nato poštni opomnik za spletno anketo s poštnim vprašalnikom in najavo telefonskega anektiranja, na koncu pa bi

sledilo telefonsko anketiranje. Z ločitvijo faz kontaktiranja anketirancev in zbiranja podatkov bi tako ocenili fiksne in variabilne stroške posamezne faze ter izračunali stopnjo optimalnosti kot produkt MSE in stroškov. Z opazovanjem sprememb med posameznimi fazami izvedbe ankete bi ugotovili, v kolikšni meri uvajanje dodatnih (dražjih) načinov prispeva k optimalnosti med kakovostjo podatkov in stroški raziskovanja.

7 Literatura

1. Belak, Eva. 2007. *Vpliv uporabe mobilnih telefonov na anketno zbiranje podatkov*. Magistrsko delo. Ljubljana: Fakulteta za družbene vede.
2. Belak, Eva in Polona Zdešar. 2008. *Opis anketiranja in analiza neodgovorov pri Anketi o uporabi informacijsko-komunikacijske tehnologije v gospodinjstvih in pri posameznikih*. Ljubljana: Statistični urad Republike Slovenije.
3. Belak, Eva in Martina Kontelj. 2008. *Standardno poročilo o kakovosti za raziskovanje: Anketa o žrtvah kriminala*. Ljubljana: Statistični urad Republike Slovenije.
4. Biemer, P. Paul. 2010. Total Survey Error: Design, Implementation, and Evaluation. *Public Opinion Quarterly* 74 (5). Dostopno prek: <http://poq.oxfordjournals.org/content/74/5/817.full.pdf> (9. julij 2011).
5. --- in Lars E. Lyberg. 2003. *Introduction to survey quality*. Hoboken, NJ: Wiley.
6. Callegaro, M. 2004. *Challenges in measuring »well-being« in a timely manner over the phone*. Dostopno prek: <http://www.gallup-europe.be/archives/events/meee04/Presentations/Mario%20Callegaro%20MEEE%2004.pdf> (15. junij 2011).
7. Cochran, G. William. 1978. *Sampling techniques (3rd ed.)*. New York: John Wiley.
8. Czaja, R. in J. Blair. 2004. *Designing surveys: A guide to decisions and procedures*. Dostopno prek: <http://www.uk.sagepub.com/books/Book225631/> (27. junij 2011).
9. Couper, M. 2002. *New Technologies and Survey Data Collection: Challenges and Opportunities*. Dostopno prek: <http://www.websm.org/uploadi/editor/new%20tehnologies%20and%20survey%20data%20collection.pdf> (2. maj 2011).
10. de Leeuw, E. D. 2005. To mix or not to mix data collection modes in surveys. *Journal of Official Statistics* (21): 233–255.

11. Deming, W. Edwards. 1950. *Some theory of sampling*. New York: Wiley.
12. Dillman, Don. 2000. *Mail and Telephone Surveys: The Tailored Design Method*. New York: Wiley.
13. ESOMAR. 2005. *Guideline on Conducting Market and Opinion Research using the Internet*. Dostopno prek: <http://www.esomar.org/web/show/id=49859> (7. avgust 2011).
14. Evropska Komisija. 2009. *Commission Staff Working Paper: Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions—Europe's Digital Competitiveness Report 2009*. Dostopno prek: http://ec.europa.eu/information_society/eeurope/i2010/docs/annual_report/2009/sec_2009_1103.pdf (28. junij 2011).
15. Gaziano, Cecilie. 2005. Comparative Analysis of within- household respondent selection Techniques. *Public Opinion Quarterly* 69 (1). Dostopno prek: https://www.webdepot.umontreal.ca/Enseignement/SOCIO/Intranet/Sondage/public/exemples_public/Gaziano_Comparative_Analysis.pdf (27. junij 2011).
16. Groves, M. Robert. 1989. *Survey Errors and Survey Costs*. New York: Wiley.
17. --- 2005. *Total survey error: Past, present, and future*. Dostopno prek: <http://poq.oxfordjournals.org/content/74/5/849.full> (20. avgust 2011)
18. --- in Robert D. Tortora. 1998. *Integrating CASIC into existing designs and organisations: a survey of the field*. John Wiley & Sons.
19. Groves, M. Robert, J. F. Fowler, M. P. Couper, J. M. Lepkowski, E. Singer in R. Tourangeau. 2004. *Survey methodology*. Hoboken, NJ: Wiley.

20. Groves, M. Robert, S. Presser in S. Dipko. 2004. The role of topic interest in survey participation decisions. *Public Opinion Quarterly* 68 (1). Dostopno prek: <http://poq.oxfordjournals.org/content/68/1/2.full.pdf> (28. junij 2011).
21. Kalton, G. in Vasja Vehovar. 2001. *Vzorčenje v anketah*. Ljubljana: Fakulteta za družbene vede.
22. Kish, Leslie. 1965. *Survey sampling*. New York: Wiley.
23. Jowell, R.. 1998. How comperative is comperative research?. *American Behavioral Scientist* (42) : 168-177. Dostopno prek: http://www.soc.aau.dk/fileadmin/user_upload/kbm/VoF/Kurser/2010/Research-design/litterature/Jowell-Roger-Comparative-research.pdf (13. julij 2011).
24. Lessler, J. T. in Kalsbeek, W. D. 1992. *Nonsampling error in surveys*. New York: Wiley.
25. Lozar Manfreda, Katja. 2001. *Web survey errors*. Doktorska disertacija. Ljubljana: Fakulteta za družbene vede.
26. ---, Vasja Vehovar in Zanel Batagelj. 2000. Veljavnost interneta kot anketnega orodja. *Teorija in praksa* 37 (6). Dostopno prek: <http://dk.fdv.uni-lj.si/tip/tip20006manfreda-vehovar.PDF> (28. junij 2011).
27. --- 2001. *Web versus Mail Questionnaire for an Institutional Survey*. In A. Westlake, The Challenge of the Internet Association for Survey Computing. London: Association for survey computing.
28. Lozar Manfreda, Katja, Nejc Berzelak in Vasja Vehovar. 2009. *Can Web be an alternative to expensive faceto-face/telephone surveys of the general population? Comparison of survey errors & costs*. Paper presented at the 3rd Workshop on Measurement and Experimentation with Internet Panels, Santpoort, the Netherlands.

29. Nathan, G. 2001. Telesurvey Methodologies for Household Surveys – A Review and Some Thoughts for the Future. *In Survey Methodology* 27 (1). Dostopno prek: <http://pluto.mscc.huji.ac.il/~gad/papers/Nathan%20SM%202001.pdf> (2. maj 2011).
30. Statistični urad RS. Dostopno prek: <http://www.stat.si> (20. avgust 2011).
31. --- 2003. *Standardno Poročilo o Kakovosti: Priporočilo, sprejeto na 5. sestanku delovne skupine za ocenjevanje kakovosti v statistiki v Luxembourggu 2002*. Ljubljana: Statistični urad Republike Slovenije
32. Vehovar, Vasja in Eva Belak. 2004. *Vpliv uporabe mobilnih telefonov na anketno zbiranje podatkov*. Dostopno prek: www.stat.si/radenci/program_2004/c6-vehovar,belak.doc (26. junij 2011).
33. Vehovar, Vasja in Katja Lozar Manfreda. 2008. Overview: Online Surveys. V *The SAGE Handbook of Online Research Methods*, ur. N. Fielding, R. M. Lee in G. Blank, 177–195. London: SAGE Publications.
34. Vehovar, Vasja, Nejc Berzelak in Katja Lozar Manfreda. 2010. Mobile phones in an Environment of Competing Survey Modes: Applying Metric for Evaluation of Costs and Errors. *Social Science Computer Review* 28 (3): 303–317.
35. Voogt, R. J. J. in Saris, W. E.. 2005. Mixed mode designs: Finding the balance between nonresponse bias and mode effects. *Journal of Official Statistics* 21 (3): 367–387.
36. Zdešar, Polona. 2008. *Uporaba informacijsko-komunikacijske tehnologije v gospodinjstvih in po posameznikih 2008*. Ljubljana: Statistični urad Republike Slovenije.

Prilogi

Priloga A: Tabele ocenitve stroškov posameznega načina

Tabela A.1: Ocenitev administrativnih stroškov za telefonsko in terensko anketiranje

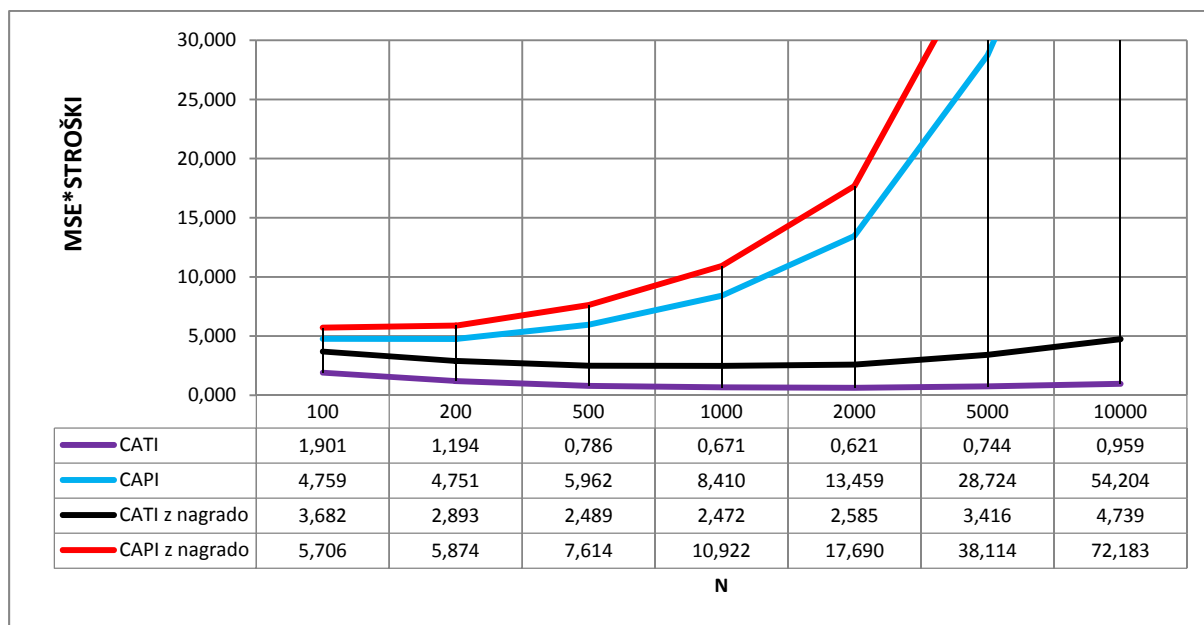
Naziv	Fiks/Var	Variabilnost glede na	Enota	Cena enote
Vzorčenje				
Izbor vzorca	Fiks	/	kos	100,00 €
Telefonsko anketiranje (CATI, M-CATI)				
Programiranje telefonskega vprašalnika	Fiks	/	ura	20,00 €
Iskanje telefonskih števil izbranih enot	Var	Nerespondenti predhodne faze	ura	6,70 €
Klic na fiksni telefon	Var	Nerespondenti predhodne faze	minuta	0,03 €
Klic na mobilni telefon	Var	Nerespondenti predhodne faze	minuta	0,10 €
Delo anketarja	Var	Nerespondenti predhodne faze	ura	5,00 €
Spremljanje zbiranja podatkov	Var	???	ura	20,00 €
Čiščenje podatkov in priprava	Var	Respondenti faze	ura	20,00 €
Terensko anketiranje				
Postavitev in priprava vprašalnika za terensko anketiranje	Fiks	/	ura	20,00 €
List A4	Var	Nerespondenti predhodne faze	kos	0,01 €
Tisk	Var	Nerespondenti predhodne faze	kos	0,03 €
Delo anketarja (izpolnjena anketa)	Var	Respondenti faze	kos	10,00 €
Delo anketarja (neizpolnjena anketa)	Var	Nerespondenti faze	kos	3,00 €
Potni stroški anketarja	Var		km	0,37 €
Spremljanje zbiranja podatkov	Var		ura	20,00 €
Vnos podatkov	Var	Respondenti faze	ura	6,70 €
Čiščenje podatkov in priprava	Var	Respondenti faze	ura	20,00 €

Tabela A.2: Parametri telefonske in terenske ankete

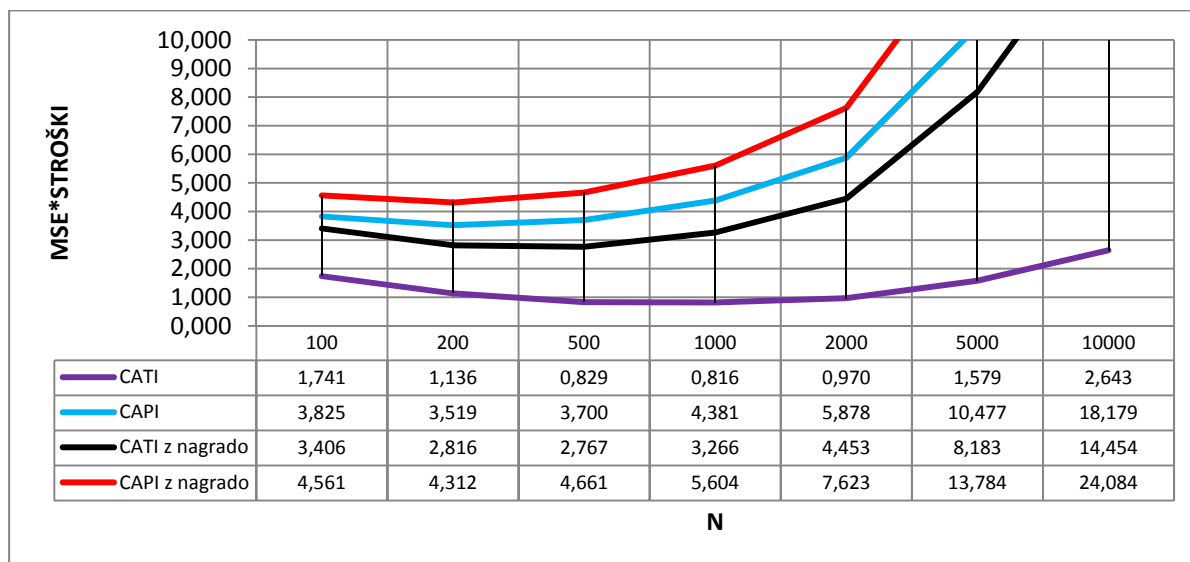
Telefonsko anketiranje				
Čas za programiranje vprašalnika	10,00	ur		
Povprečen čas za iskanje ene številke	1,20	minut	0,02	ur
Povprečen čas za identifikacijo neustrezne RDD-številke		minut	0,00	ur
Povprečen čas rekrutacije	1,00	minut	0,02	ur
Povprečen čas celotne ankete	9,80	minut	8,80	minut brez rekrutacije 0,15 ur brez rekrutacije
Povprečen čas ankete z menjavo načina	2,19	minut	1,19	minut brez rekrutacije 0,02 ur brez rekrutacije
Povprečno število poskusov rekrutacije	5,00	poskusov, vključno z uspešnim		
Povprečni čas spremljanja zbiranja podatkov na ????	0,00	ur		
Povprečni čas čiščenja in priprave podatkov na respondenta	0,50	minut	0,01	ur
Terensko anketiranje				
Čas za postavitev in pripravo vprašalnika	15,00	ur		
Število strani poštnega vprašalnika	16,00	strani		
Število listov poštnega vprašalnika	8,00	listov		
Povprečno število kilometrov na enoto iz vzorca z upoštevanjem vseh kontaktov	7,00	kilometrov		
Povprečni čas spremljanja zbiranja podatkov na ????	0,00	ur		
Povprečni čas vnosa enega vprašalnika	6,00	minut	0,10	ur
Povprečni čas čiščenja in priprave podatkov na respondenta	2,00	minut	0,03	ur

Priloga B: Simulacije spremenljivk

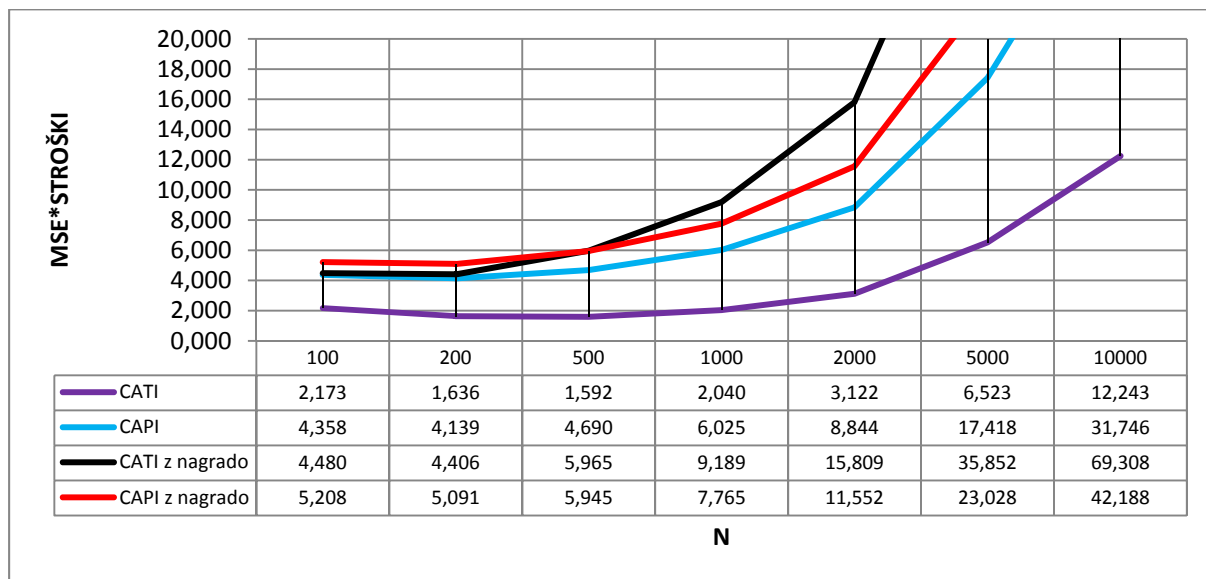
Slika B.1: Simulacija spremenljivke »uporaba interneta« brez nagrade in z nagrado



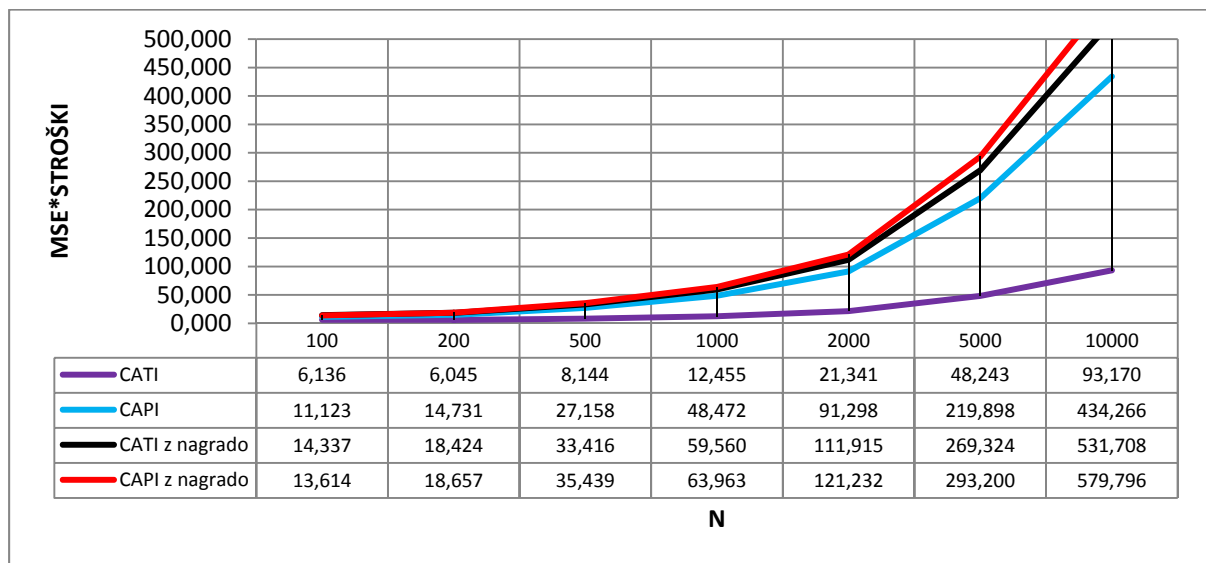
Slika B.2: Simulacija spremenljivke »uporaba računalnika« brez nagrade in z nagrado



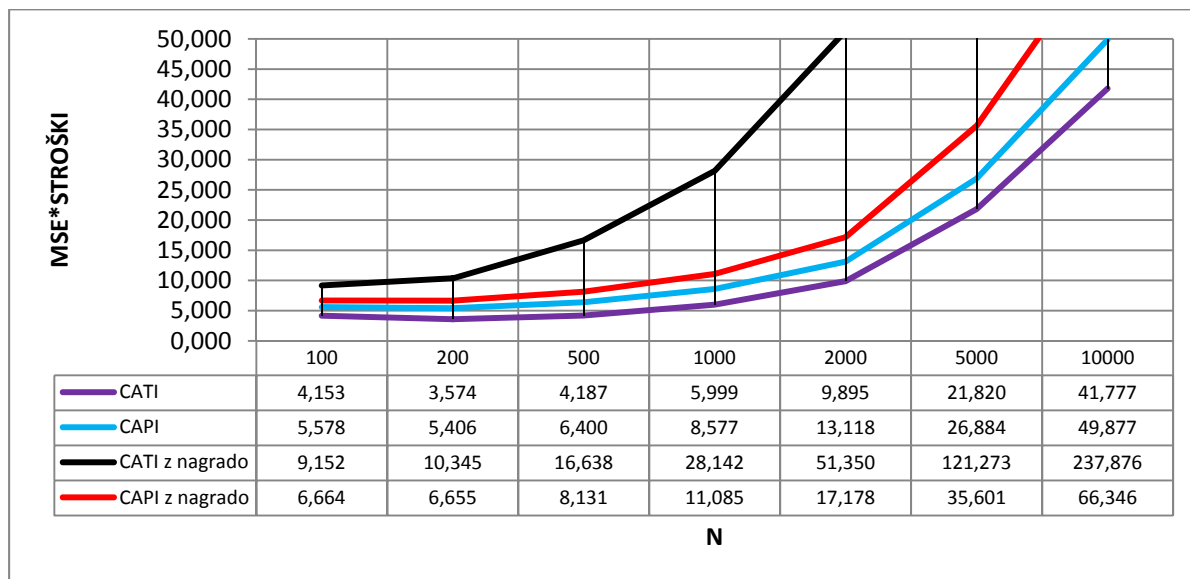
Slika B.3: Simulacija spremenljivke »pogostost uporabe računalnika« brez nagrade in z nagrado



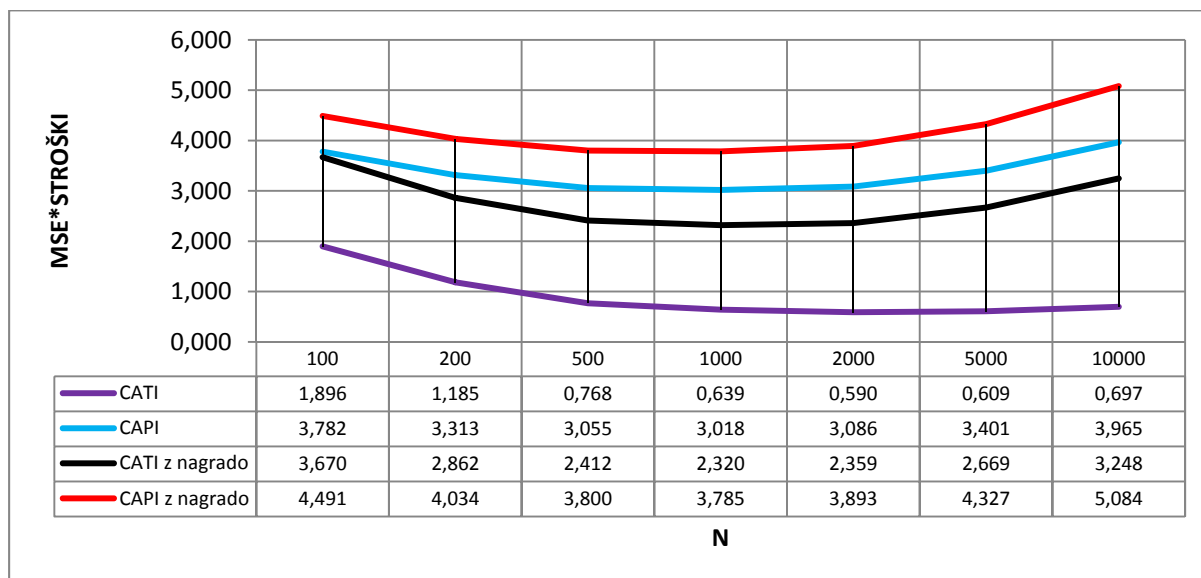
Slika B.4: Simulacija spremenljivke »plačljive vsebine so kakovostnejše od brezplačnih« brez nagrade in z nagrado



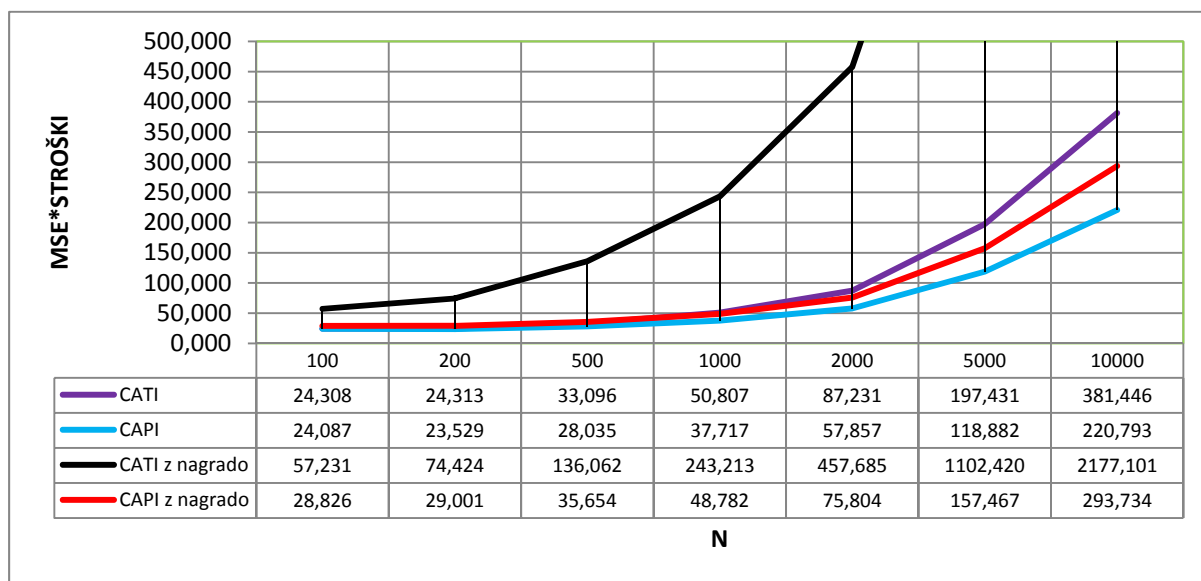
Slika B.5: Simulacija spremenljivke »nadomestitev kupovanja s prenašanjem vsebin iz interneta« brez nagrade in z nagrado



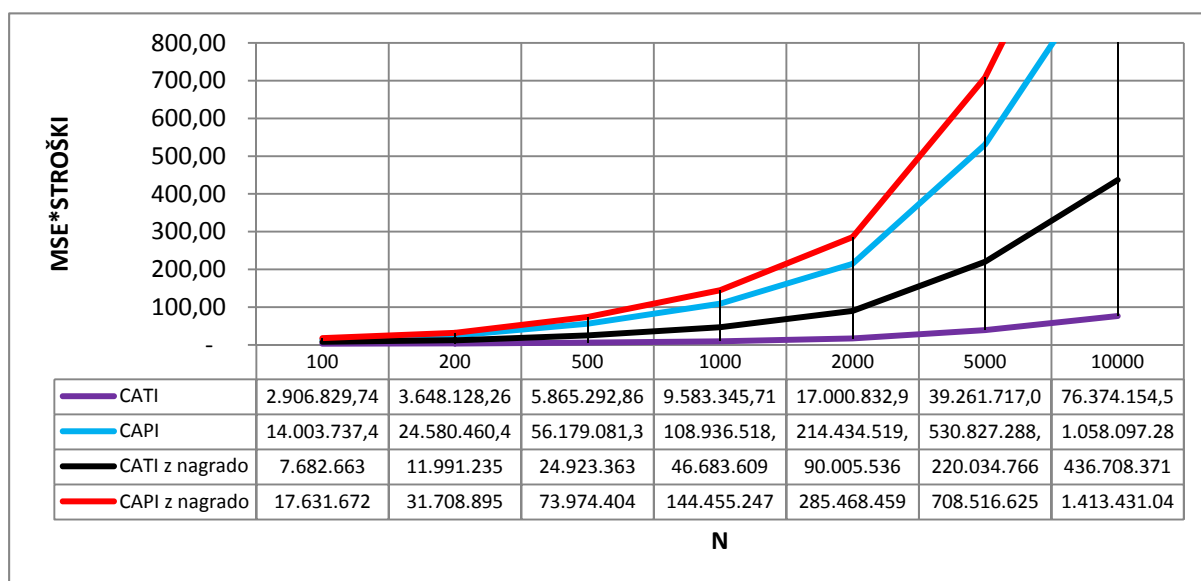
Slika B.6: Simulacija spremenljivke »izobrazba« brez nagrade in z nagrado



Slika B.7: Simulacija spremenljivke »število IKT-naprav« brez nagrade in z nagrado



Slika B.8: Simulacija spremenljivke »dohodek« brez nagrade in z nagrado



Slika B.9: Simulacija spremenljivke »starost« brez nagrade in z nagrado

