

UNIVERZA V LJUBLJANI
FAKULTETA ZA DRUŽBENE VEDE

Pia Pirc Žagar

Priistranskost zaradi nepokritja v spletnih anketah

Magistrsko delo

Ljubljana, 2017

UNIVERZA V LJUBLJANI
FAKULTETA ZA DRUŽBENE VEDE

Pia Pirc Žagar

Mentorica: izr. prof. dr. Katja Lozar Manfreda

Pristranskost zaradi nepokritja v spletnih anketah

Magistrsko delo

Ljubljana, 2017

Zahvala

Rada bi se zahvalila mentorici dr. Katji Lozar Manfreda za pomoč in nasvete pri pisanju tega magistrskega dela ter bližnjim za podporo.

Pristranskost zaradi nepokritja v spletnih anketah

Eden od glavnih virov napak v spletnih anketah je pristranskost zaradi nepokritja, do katere pride, ker v vzorec lahko zajamemo le uporabnike interneta. Pristranskost zaradi nepokritja je še posebej visoka, kadar je v preučevani populaciji velik delež neuporabnikov interneta ali pa so med uporabniki in neuporabniki velike razlike. Ob vse višji penetraciji interneta lahko pričakujemo čedalje nižje vrednosti pristranskosti zaradi nepokritja, kar prinaša potencial za izvajanje verjetnostnih spletnih anket splošne populacije. Na ta način bi se lahko izognili dragim metodam za vključevanje neuporabnikov interneta, kot so verjetnostni spletni paneli, kjer neuporabnikom priskrbijo ustrezno tehnično opremo. V tem magistrskem delu obravnavamo spremembe v ocenah pristranskosti zaradi nepokritja v potencialnih spletnih anketah za obdobje med leti 2010 in 2016. S tem nadaljujemo analizo pristranskosti zaradi nepokritja, ki so jo izvedli Mohorko in drugi (2011) za leta 2005 do 2009. Analiza temelji na podatkih iz Evrobarometra, ki so bili zbrani preko terenskih anket in omogočajo primerjavo med uporabniki in neuporabniki interneta za pet ključnih spremenljivk: spol, starost, starost ob koncu rednega izobraževanja, politična usmeritev levo-desno in zadovoljstvo z življenjem. Rezultati kažejo, da bi ob izključitvi neuporabnikov podcenili delež žensk, povprečno starost in desno politično usmeritev ter precenili starost pri končani izobrazbi in zadovoljstvo z življenjem. V obdobju od leta 2010 do leta 2016 se je pristranskost v primerjavi z obdobjem od 2005 do 2009 zmanjšala za vse spremenljivke, vendar se je proces upadanja z leti nekoliko upočasnil. Pri nizkih vrednostih pristranskosti namesto nadaljnjega upadanja opažamo oscilacije.

Ključne besede: spletne ankete, pristranskost anketnih ocen, napaka pokritja, pristranskost zaradi nepokritja, Evrobarometer.

Bias Due to Non-Coverage in Web Surveys

One of the main sources of survey error affecting web surveys is bias due to non-coverage, which is caused by the fact that their samples are limited to internet users. This bias is particularly high when the studied population includes a large share of non-users or when there are large differences between users and non-users. With the growth of internet penetration, we expect the bias due to non-coverage to decrease, which is a promising sign for web-based probability surveys of the general population. Such surveys could replace expensive probability-based online panels, which provide technical equipment for non-users. This master thesis explores the changes in bias of estimates due to non-coverage for web surveys for the period between 2010 and 2016. It follows up on a previous study by Mohorko et al. (2011), which traced coverage bias between 2005 and 2009. The analysis uses Eurobarometer data collected with face-to-face surveys, allowing us to compare internet users and non-users on five key variables: gender, age, terminal education age, political left-right self-placement and general life satisfaction. The results show that excluding internet non-users would result in underestimating the share of women, the average age and right-wing political leanings. It would also result in overestimating the terminal education age and life satisfaction. The bias due to non-coverage in the analysed period has decreased in comparison to 2005 to 2009 for all variables, but the rate of the decrease has somewhat slowed down through the years. With low values of bias, we can notice oscillations instead of a consistent downward trend.

Keywords: web surveys, survey bias, coverage error, bias due to non-coverage, Eurobarometer.

Kazalo vsebine

1	UVOD	8
2	SPLETNE ANKETE: PREDSTAVITEV IN PROBLEM REPREZENTATIVNOSTI.....	11
3	DIGITALNI RAZKORAK	15
3.1	Penetracija interneta	15
3.2	Razlike med posamezniki na nasprotnih straneh digitalnega razkoraka	17
4	TEORIJA ANKETNIH NAPAK IN NAPAKE POKRITJA.....	21
4.1	Opredelitev pojma anketne napake	21
4.2	Napaka pokritja	22
5	MERJENJE PRISTRANSKOSTI ZARADI NEPOKRITJA V SPLETNIH ANKETAH	27
5.1	Primer panela Health and Retirement Study v ZDA	27
5.2	Primer mobilnih spletnih anket v Evropi	29
5.3	Primer panela LISS	30
5.3.1	Leenheer in Scherpenzeel (2013)	30
5.3.2	Eckman (2015).....	33
5.3.3	Toepoel in Hendriks (2016).....	35
6	EMPIRIČNA ANALIZA PODATKOV IZ EVROBAROMETRA	37
6.1	Raziskovalno vprašanje	38
6.2	Podatki in uteževanje	38
6.3	Metodologija.....	40
6.4	Analiza in rezultati	42

6.4.1	Penetracija interneta	42
6.4.2	Razlike med uporabniki in neuporabniki interneta.....	45
6.4.3	Absolutna pristranskost zaradi nepokritja	49
6.4.4	Relativna pristranskost in primerjava z leti 2005 do 2009	53
6.4.5	Analiza relativne pristranskosti po državah	56
7	SKLEP.....	66
8	LITERATURA.....	70
	PRILOGE	73
	Priloga A: Uporabljena anketna vprašanja in iz njih izpeljane spremenljivke	73
	Priloga B: Penetracija interneta v gospodinjstvih za obdobje 2010–2016.....	74
	Priloga C: Relativna in absolutna pristranskost zaradi nepokritja po letih in državah	75
 Kazalo slik		
	Slika 3.1: Odstotek gospodinjstev z dostopom do interneta	16
	Slika 3.2: Odstotek posameznikov, ki so v zadnjih 12 mesecih uporabljali internet.....	17
	Slika 3.3: Odstotek posameznikov, ki so v zadnjih 12 mesecih uporabljali internet, po starostnih skupinah za leto 2016	18
	Slika 3.4: Odstotek posameznikov v Sloveniji, ki so v zadnjih 12 mesecih uporabljali internet, glede na spol in starostno skupino za leto 2016.....	18
	Slika 3.5: Odstotek posameznikov v EU, ki so v zadnjih 12 mesecih uporabljali internet, glede na spol in starostno skupino za leto 2016.....	19
	Slika 3.6: Odstotek posameznikov v Sloveniji in EU, ki so v zadnjih 12 mesecih uporabljali internet, glede na izobrazbo	20
	Slika 4.1: Shema problema nepokritja	23
	Slika 6.1 Rast dostopa do interneta od doma med leti 2005 in 2016	44
	Slika 6.2 Absolutna pristranskost zaradi nepokritja pri spremenljivki spol	50
	Slika 6.3: Absolutna pristranskost zaradi nepokritja pri spremenljivki starost	51
	Slika 6.4 Absolutna pristranskost zaradi nepokritja pri spremenljivki starost pri končani izobrazbi.....	52

Slika 6.5 Absolutna pristranskost zaradi nepokritja pri spremenljivki politična usmeritev	52
Slika 6.6 Absolutna pristranskost zaradi nepokritja pri spremenljivki zadovoljstvo z življenjem.....	53
Slika 6.7: Relativna pristranskost za vse spremenljivke po letih.....	55
Slika 6.8: Vrednosti relativne pristranskosti za Slovenijo	55
Slika 6.9: Povprečna (absolutna) relativna pristranskost za leto 2016 za analizirane države ..	58
Slika 6.10: Povprečna (absolutna) relativna pristranskost za leta 2010 do 2016 za analizirane države	59
Slika 6.11: Relativna pristranskost za spol skozi čas	61
Slika 6.12: Relativna pristranskost za starost skozi čas	62
Slika 6.13: Relativna pristranskost za starost pri dokončani izobrazbi skozi čas	63
Slika 6.14: Relativna pristranskost za politično usmeritev skozi čas	64
Slika 6.15: Relativna pristranskost za zadovoljstvo z življenjem skozi čas	65

Kazalo tabel

Tabela 6.1 Primerjave relativne pristranskosti za leto 2011 glede na uporabljeno utež.....	40
Tabela 6.2: Frekvence za spremenljivko <i>anketiranec ima doma internetno povezavo</i>	42
Tabela 6.3: Bivariatne analize	45
Tabela 6.4: Bivariatne analize za Slovenijo za leto 2010	47
Tabela 6.5: Absolutna pristranskost po letih za vse spremenljivke	50
Tabela 6.6: Relativna pristranskost po letih.....	54
Tabela 6.7: Države glede na povprečno relativno pristranskost*	57
Tabela 6.8: Države po skupinah glede na stopnjo penetracije interneta leta 2016	59
Tabela 6.9: Povprečna relativna pristranskost za leta 2010 do 2016 po skupinah držav glede na stopnjo penetracije interneta leta 2016.....	60

1 Uvod

Danes lahko vpliv interneta čutimo na vsakem koraku. Razvoj informacijsko-komunikacijskih tehnologij je preobrazil tudi področje anketiranja, kjer zdaj prevladujejo spletne ankete. Le te so se od svojih začetkov pred več kot dvema desetletjema uveljavile kot cenejša in hitrejša alternativa za starejše oblike anketiranja, kot so telefonske in terenske ankete. Upravičeno torej pričakujemo, da bo tudi nadaljnji razvoj na tem področju tesno povezan z napredkom novih tehnologij. S tem ko se vse večji delež vseh anket izvaja preko spleta, pa postaja čedalje bolj pomembno, da se zavedamo specifičnih anketnih napak, ki vplivajo na rezultate spletnih anket, in iščemo načine za izboljšanje kakovosti podatkov iz takšnih anket.

V tem magistrskem delu se osredotočamo na eno izmed glavnih napak v spletnih anketah – napako pokritja. Do te napake pride, ker v spletni vzorec ne moremo zajeti članov populacije brez dostopa do interneta. To je problematično zato, ker med uporabniki in neuporabniki interneta lahko obstajajo precejšnje demografske in socioekonomske razlike. Z leti je neuporabnikov interneta čedalje manj, kar pomeni, da s spletnimi anketami lahko potencialno pokrijemo vse večji del prebivalstva. Ob tem se postavlja vprašanje, kako se rast penetracije interneta odraža v natančnosti ocen, pridobljenih s spletnimi anketami. Spremembe v pristranskosti zaradi nepokritja v Evropi za obdobje med leti 2005 in 2009 so raziskali Mohorko in drugi (2011), ki so izvedli analizo na podlagi podatkov iz raziskave Evrobarometer. Pokazali so, da se z leti stopnja penetracije interneta zvišuje, s tem pa se manjša tudi pristranskost zaradi nepokritja, hitrost tega procesa pa se razlikuje od države do države. Kljub temu so razlike med uporabniki in neuporabniki interneta precejšnje tudi v državah, kjer ima dostop do interneta nad 80 % prebivalcev.

Osnovni namen tega magistrskega dela je izvesti analizo pristranskosti zaradi nepokritja med leti 2010 in 2016 po zgledu Mohorko in drugih (2011), da bi videli, kako so se ocene pristranskosti zaradi nepokritja spreminjale v zadnjih letih. Glavno raziskovalno vprašanje, ki si ga zastavljamo, je, kakšen bi bil vpliv na rezultate anket, če iz vzorca izključimo uporabnike interneta. Pričakovati gre, da je v zadnjih sedmih letih pristranskost zaradi nepokritja ob vse večji penetraciji interneta še bolj upadla, ob tem pa se poraja tudi vprašanje, ali bi spletne ankete lahko že uporabljali za verjetnostno anketiranje splošne populacije. Do zdaj je bilo za verjetnostne vzorce splošne populacije potrebno kombiniranje anketnih

načinov (npr. anketiranje neuporabnikov interneta po telefonu, medtem ko ostali anketiranci izpolnjujejo spletno obliko ankete) ali oblikovanje verjetnostnih spletnih panelov, pri katerih vzorčenje izvajajo na splošni populaciji, nato pa neuporabnikom interneta priskrbijo opremo za sodelovanje v spletnih anketah. Takšni pristopi (zlasti verjetnostni spletni paneli) zahtevajo precej večje finančne investicije od običajnih spletnih anket, zato nas zanima, ali še vedno lahko trdimo, da pomembno izboljšajo kakovost podatkov. Z drugimi besedami, postavlja se vprašanje, ali jih zares še vedno potrebujemo ali pa bi se lahko osredotočili le na uporabnike interneta, tako kot so se v 90-ih letih prejšnjega stoletja anketni raziskovalci z množično uporabo telefonskih anket osredotočili le na tiste, ki so telefon imeli.

Da bi odgovorili na zgornja vprašanja, bomo izvedli analizo na podlagi podatkov iz raziskave Evrobarometer, ki so bili zbrani s pomočjo terenskih anket in vključujejo spremenljivko, ki meri dostop do interneta. S pomočjo te spremenljivke bomo anketirance razdelili na uporabnike in neuporabnike interneta in ocenili, kakšna bi bila pristranskost zaradi nepokritja, če bi bili neuporabniki izključeni iz raziskave, kar se zgodi pri spletnih anketah. Analizo bomo izvedli na uteženih podatkih za vzorec, ki, tako kot pri Mohorko in drugih (2011), vsebuje vse države članice Evropske unije, Turčijo in turško skupnost na Cipru (angl. *Turkish Cypriot Community*, v nadaljevanju TCC). Poleg absolutne in relativne pristranskosti zaradi nepokritja, bomo predstavili tudi rezultate bivariatnih statističnih analiz razlik med uporabniki in neuporabniki interneta.

Teoretični del magistrskega dela je razdeljen na štiri poglavja. V prvem poglavju predstavimo metodo spletnega anketiranja, nato pa se osredotočimo na potencial verjetnostnih spletnih anket za raziskovanje splošne populacije in si ogledamo trenutno stanje na tem področju. V naslednjem poglavju se ukvarjamo z digitalnim razkorakom med uporabniki in neuporabniki interneta, ki predstavlja eno od glavnih ovir pri reprezentativnosti rezultatov spletnih anket. Prikažemo podatke o penetraciji interneta v Evropski uniji in si ogledamo razlike med uporabniki in neuporabniki interneta pri nekaterih demografskih spremenljivkah. Tretje poglavje je namenjeno teoriji anketnih napak in opredelitvi napake pokritja ter pristranskosti zaradi nepokritja. Tukaj predstavimo tudi formule, ki jih bomo uporabili za račun pristranskosti v empiričnem delu. Zadnje poglavje teoretičnega dela zajema pregled nekaterih študij, ki so se ukvarjale s pristranskostjo zaradi nepokritja v spletnih anketah. Ogledamo si metodologijo teh študij ter njihove rezultate. Sledi empirični del, kjer najprej predstavimo študijo Mohorko in drugih (2011), na kateri temelji naša analiza, nato pa nadaljujemo s

predstavitvijo raziskovalnih vprašanj, podatkov in metodologije. Rezultati analize podatkov iz Evrobarometra so razdeljeni na pet delov. Najprej predstavimo podatke o deležu uporabnikov interneta, temu pa sledijo primerjava uporabnikov in neuporabnikov s pomočjo bivariatnih analiz, analiza pristranskosti zaradi nepokritja, pregled sprememb pristranskosti po letih v primerjavi z rezultati Mohorko in drugih (2011) ter pregled razlik v pristranskosti med državami v vzorcu.

2 Spletne ankete: predstavitev in problem reprezentativnosti

V tem poglavju predstavimo metodo spletnega anketiranja in njene prednosti ter slabosti, nato pa se osredotočimo na verjetnostne spletne ankete, saj le te omogočajo statistično sklepanje iz vzorca na populacijo. Posebej nas zanima potencial verjetnostnih spletnih anket za raziskovanje splošne populacije. Ogledamo si, kaj so ovire na tem področju, ter predstavimo pristope, ki se trenutno uporabljajo za njihovo premagovanje.

Anketa je metoda sistematičnega zbiranja podatkov, pri kateri anketiranci odgovarjajo na standardiziran vprašalnik. Njen končni namen je kvantitativna analiza neke ciljne populacije, iz katere smo zajeli vzorec vprašanih. Tradicionalne metode anketiranja vključujejo terensko anketiranje (angl. *face to face*) in ankete po pošti, kasneje pa tudi ankete prek telefona (Callegaro in drugi 2015, 4). Internet se je kot orodje za izvajanje anket začel uporabljati konec osemdesetih let prejšnjega stoletja, ko se je uveljavilo anketiranje preko e-pošte, začetki interaktivnih spletnih anket, kot jih poznamo danes, pa segajo v sredino devetdesetih let (Lozar Manfreda in Vehovar 2009, 264). Ključne značilnosti takšnih anket so, da so računalniško podprte, da se nahajajo na strežniku, kamor se shranijo tudi odgovori, in da anketiranci do njih dostopajo preko spletnega brskalnika in jih izpolnjujejo sami brez prisotnosti raziskovalca (Callegaro in drugi 2015, 5).

Danes so spletne ankete zelo dobro uveljavljene v komercialnem sektorju (npr. za tržno raziskovanje), pa tudi pri akademskem raziskovanju in uradnih statistikah v vladnem sektorju, čeprav le ta nekoliko zaostaja za trgovino. Zaradi številnih prednosti, ki jih naštevamo v nadaljevanju, so se uveljavile tudi pri tako imenovanem »naredi sam« raziskovanju (angl. *DIY research*), ki ga izvajajo posamezniki izven okvirov večjih institucij (Callegaro in drugi 2015, 6–7).

Spletne ankete so s seboj prinesle številne prednosti, kot so nizki stroški in lažji dostop do večjega števila anketirancev, ter posledično demokratizacijo raziskovanja, ker si zbiranje podatkov preko spletnih anket lahko privoščijo tudi posamezniki in manjše organizacije (Couper 2000, 464–465). Poleg tega je spletno anketiranje veliko hitrejšo od klasičnih pristopov, kar prispeva k učinkovitosti, ne pa nujno tudi h kakovosti raziskovanja (Couper in Miller 2007, 832). Nove tehnologije so spremenile tudi tehnične vidike anket, saj ponujajo možnosti, kot so vključitev multimedijskih vsebin in programiranje pogojev, s katerimi lahko

določimo, katerim anketirancem bo zastavljeno določeno vprašanje. Še ena prednost je občutek zasebnosti, zaradi katerega je lažje pridobiti odgovore, ki se nanašajo na občutljive teme (de Leeuw 2012, 75).

Navkljub omenjenim prednostim in veliki priljubljenosti spletnega anketiranja obstajajo tudi skrbi glede metodološke ustreznosti spletnih anket (Couper in Miller 2007, 833), na katere se opozarja že od samih začetkov tega pristopa (gl. Couper 2000). Couper (2000, 466) poudarja, da na kakovost anket ne moremo gledati kot na nekaj absolutnega, temveč jo lahko ocenimo le, če upoštevamo ostale faktorje anketnega načrta, kot so cena, natančnost in časovni razpon, in če naredimo primerjavo z drugimi možnimi pristopi, ki bi jih lahko izbrali. Glavni problem so še vedno napake neopazovanja (angl. *errors of nonobservation*, gl. četrto poglavje) in problem reprezentativnosti (Couper in Miller 2007, 833), ki je tesno povezan s problemom vzorčenja.

Pri vseh vrstah anket poznamo dva načina vzorčenja: verjetnostno vzorčenje (angl. *probability sampling*) in neverjetnostno vzorčenje (angl. *non-probability sampling*). Samo verjetnostni vzorci omogočajo statistično sklepanje, saj ima vsaka enota v vzorčnem okviru znano neničelno možnost, da je naključno izbrana v vzorec (Callegaro in drugi 2015, 7). V nadaljevanju se bomo osredotočili na verjetnostne spletne ankete, ki temeljijo na vnaprej pridobljenem seznamu enot iz ciljne populacije¹. Verjetnostne spletne ankete na podlagi seznama (angl. *list-based probability surveys*) so lahko ankete specifičnih populacij, ankete splošne populacije in verjetnostni spletni paneli. Za statistično sklepanje je ključno, da vzorčni okvir ustrezno pokrije ciljno populacijo in da so enote v vzorec izbrane s primerno verjetnostno metodo.

Primer ankete specifične populacije je anketa zaposlenih v nekem podjetju, ki imajo vsi dostop do interneta, saj v tem primeru lahko ustrezen vzorec lahko naključno izberemo s seznamom zaposlenih (Callegaro in drugi 2015, 8). Spletne ankete se zelo dobro obnesejo pri specifičnih ciljnih populacijah, pri katerih je stopnja pokritosti z internetom blizu 100 %, raziskovalci pa imajo na voljo dober vzorčni okvir, ki vsebuje e-poštna naslove vseh članov

¹ Obstajajo tudi verjetnostne spletne ankete, ki ne temeljijo na seznamu, vendar le te ne spadajo v domet tega magistrskega dela. To so prestrezne spletne ankete (angl. *intercept surveys*), zanje pa veljajo določne omejitve. Za več informacij gl. Callegaro in drugi (2015, 8).

populacije. V tem primeru je mogoče izvesti verjetnostno spletno anketo bodisi na verjetnostnem vzorcu bodisi na celotni populaciji (Callegaro in drugi 2015, 25).

Če želimo izvesti verjetnostno spletno anketo splošne populacije (angl. *general population survey*), pa za rekrutiranje anketirancev uporabimo tradicionalne metode, kot so pošta, telefon in terensko anketiranje, saj za neposredno uporabo e-pošte ali spleta nimamo ustreznega vzorčnega okvira (Callegaro in drugi 2015, 8). V tem primeru se uporablja spletne ankete kot del kombiniranega anketnega načrta (angl. *mixed mode*) ali verjetnostne spletne panele.

Pri verjetnostnih spletnih anketah splošne populacije, ki so del kombiniranega anketnega načrta, potencialne anketirance npr. najprej kontaktirajo po pošti in jim ponudijo možnost izpolnjevanja spletne ankete. Tistim, ki se na vabilo niso odzvali, nato ponudijo tudi alternativne možnosti anketiranja, npr. izpolnjevanje papirnatega vprašalnika, ki ga prejmejo po pošti (Callegaro in drugi 2015, 27).

Pri verjetnostnih spletnih panelih (angl. *probability-based online panels*), katerih ciljna populacija je splošna populacija (to so lahko posamezniki ali gospodinjstva), pa so potencialni udeleženci kontaktirani na terenu, po telefonu ali po pošti. Gospodinjstvom ali posameznikom brez interneta se za sodelovanje v panelu lahko priskrbi internetna povezava in naprava (npr. namizni ali prenosni računalnik, tablico ali pametni telefon). Primeri verjetnostnih spletnih panelov, kjer so se odločili za to možnost, so LISS na Nizozemskem, German Internet Panel v Nemčiji in GfK (Knowledge Network) panel v ZDA. Če se raziskovalci za ta korak ne odločijo, lahko ostanejo pri kombiniranemu načinu zbiranja podatkov, kjer udeleženci panela brez dostopa do interneta ankete izpolnjujejo po pošti ali telefonu. Taka primera sta Gesis panel v Nemčiji in Gallup panel v ZDA (Callegaro in drugi 2015, 27).

Oba omenjena pristopa do neuporabnikov interneta imata svoje slabosti. Razdeljevanje naprav za dostop do interneta je drago, poleg tega pa obstaja tveganje, da že sam dostop do takšne naprave vpliva na anketirance. Lahko npr. pripelje do povečane rabe interneta tudi izven časa, ki ga porabijo za sodelovanje v spletnih anketah, kar lahko privede do sprememb v siceršnjem obnašanju, stališčih itd. (t.i. »*panel conditioning*«, gl. Lynn 2009, 9). Tudi kombinirani načini zbiranja podatkov lahko negativno vplivajo na kakovost zbranih podatkov. Callegaro in drugi (2015, 38) opozarjajo na razlike med različnimi načini

anketiranja in njihovim vplivom na zbrane podatke (t.i. učinek načina anketiranja). Znan je npr. učinek primarnosti (angl. *primacy effect*) pri spletnih anketah in učinek nasledstva (angl. *recency effect*) pri telefonskih anketah. Spletni anketiranci tako disproporcionalno pogosto izberejo prvi možni odgovor, medtem ko telefonski anketiranci pogosteje izberejo odgovor, ki so ga slišali nazadnje. Da dobimo primerljive rezultate s kombiniranim načinom zbiranja podatkov, je pri načrtovanju anket in obdelavi podatkov treba upoštevati vse takšne učinke. Prav zaradi teh problemov je v zadnjem času za raziskovalce vse bolj pomembno vprašanje, ali je možno spletne ankete uporabljati za reprezentativne študije splošne populacije. Callegaro in drugi (2015, 28) za primer navajajo projekta GenPopWeb project in WebDataNet Cost Action IS 1004, ki sta se ukvarjala s to tematiko. Za Slovenijo pa je v tem pogledu še posebej pomemben projekt CRONOS, v okviru katerega so rekrutirali udeležence Evropske družboslovne raziskave (angl. *European Social Survey*) iz Estonije, Slovenije in Združenega kraljestva za sodelovanje v pilotnem verjetnostnem spletnem panelu (Villar 2016).

Če povzamemo, pri raziskovanju splošne populacije spletne ankete še vedno niso nadomestile ostalih načinov anketiranja, kljub čedalje večji penetraciji interneta. Vzroki za težave so naslednji:

- Del splošne populacije nima dostopa do interneta.
- Uporabniki in neuporabniki interneta se razlikujejo po sociodemografskih značilnostih in po življenjskem slogu.
- Raziskovalci nimajo na voljo vzorčnega okvira e-naslovov vseh članov populacije. V nasprotju s telefonskimi številkami e-naslovov ni mogoče naključno generirati, saj ne sledijo določeni strukturi. To pomeni, da morajo še vedno posegati po tradicionalnih načinih rekrutiranja, kar zviša stroške anketiranja.
- Etična načela in regulacije v posameznih državah prepovedujejo kontaktiranje naključnih e-naslovov, kjer s prejemnikom ne obstaja predhodno razmerje (Callegaro in drugi 2015, 25).

Razlike v uporabi in dostopu do interneta (t.i. digitalni razkorak), ki so vir prvih dveh težav, so bolj podrobno obdelane v naslednjem poglavju.

3 Digitalni razkorak

Kot smo omenili, je ena od glavnih ovir pri izvajanju verjetnostnih spletnih anket splošne populacije digitalni razkorak med uporabniki in neuporabniki interneta. V tem poglavju digitalni razkorak podrobneje definiramo in si na podlagi podatkov o penetraciji interneta v Evropski uniji ogledamo, kako se odraža v praksi.

Pojem digitalni razkorak (angl. *digital divide*) se nanaša na »razlike med posamezniki, gospodinjstvi, podjetji in geografskimi območji glede možnosti dostopa do informacijskih in komunikacijskih tehnologij (IKT) ter njihove uporabe« (Dolničar in drugi 2002, 83). Za potrebe tega magistrskega dela ga ožje opredelimo kot razlike v dostopu do interneta med različnimi demografskimi skupinami (Mohorko in drugi 2011, 2), pri čemer se osredotočamo predvsem na osnovni razkorak med uporabniki in neuporabniki interneta. V državah z visoko penetracijo interneta poleg tega opažajo tudi drugi digitalni razkorak (angl. *second digital divide*), ki se odraža v razlikah med uporabniki s hitro (širokopasovno) in počasno internetno povezavo (Callegaro in drugi 2015, 27)².

3.1 Penetracija interneta

Med glavnimi dejavniki, ki vplivajo na digitalni razkorak, je stopnja penetracije interneta (angl. *internet penetration rate*). Pri merjenju stopnje penetracije interneta v specifični populaciji moramo biti pozorni na potencialne vire težav. Prva težava se pojavi pri definiciji dostopa do interneta, kjer se je med drugim treba odločiti, ali nas zanimajo gospodinjstva ali posamezniki, ali merimo možnost dostopa ali dejansko uporabo interneta ter ali nas zanima samo dostop do interneta doma ali tudi drugje (npr. v službi). Drugi problem so razlike v kakovosti med anketami, s katerimi so poskušali izmeriti stopnjo penetracije; najbolj zloglasen primer je merjenje dostopa do interneta preko spletnih anket (Couper 2002, 467–468).

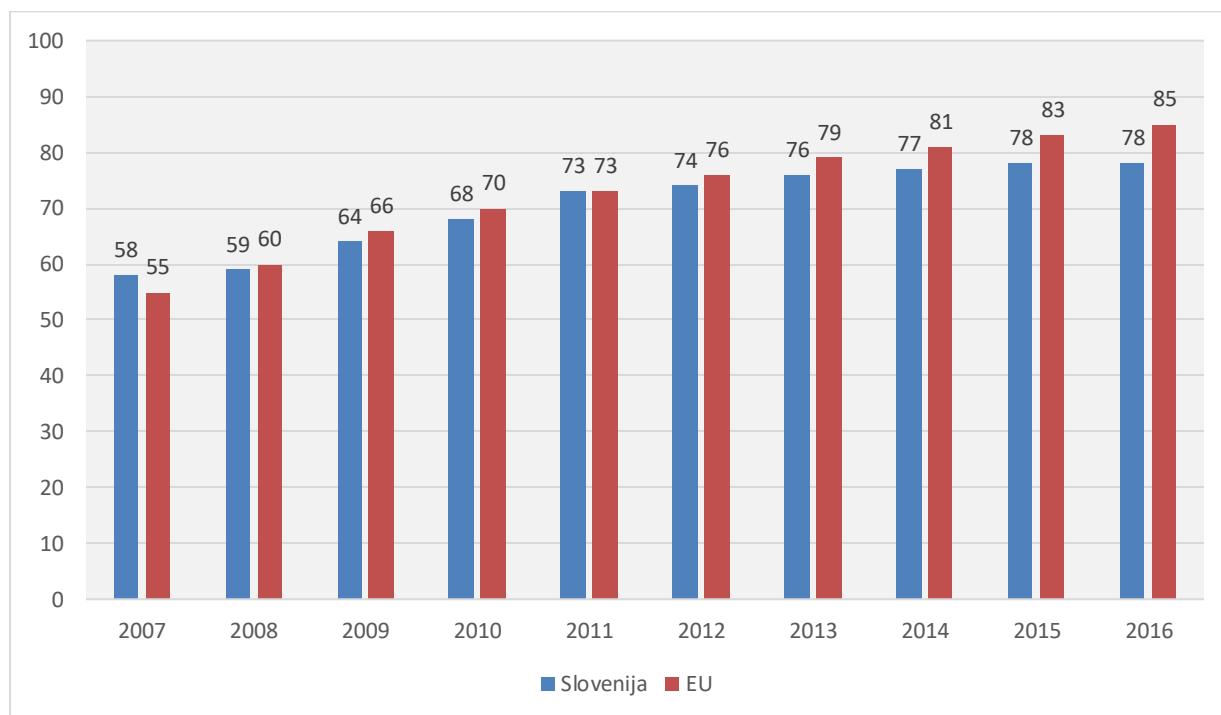
V nadaljevanju predstavljamo podatke iz Eurostatove raziskave o uporabi informacijsko-komunikacijskih tehnologij med gospodinjstvi in posamezniki, ki je bila izvedena leta 2016.

² Dolničar in drugi (2002, 87) to vrsto razkoraka imenujejo tretji digitalni razkorak, medtem ko se drugi digitalni razkorak nanaša na različne stopnje izkušenosti pri rabi interneta (angl. *experience gap*). Omenjajo tudi »dvojni« digitalni razkorak, ki se nanaša na različne skupine neuporabnikov interneta, predvsem na razlike med tistimi, ki jih internet ne zanima, in tistimi, ki sicer prepoznavajo njegovo uporabnost, vendar jih od dostopa ločijo različne ovire.

Podatki so bili zbrani osebno na terenu ali preko telefona na verjetnostnem vzorcu iz splošne populacije držav članic EU, držav kandidat, Norveške in Islandije. Dostop do interneta so merili na nivoju gospodinjstev, ki imajo vsaj enega člana v starostni skupini med 16 in 74 let, rabo interneta pa na nivoju posameznikov v tej starostni skupini (Eurostat 2017c).

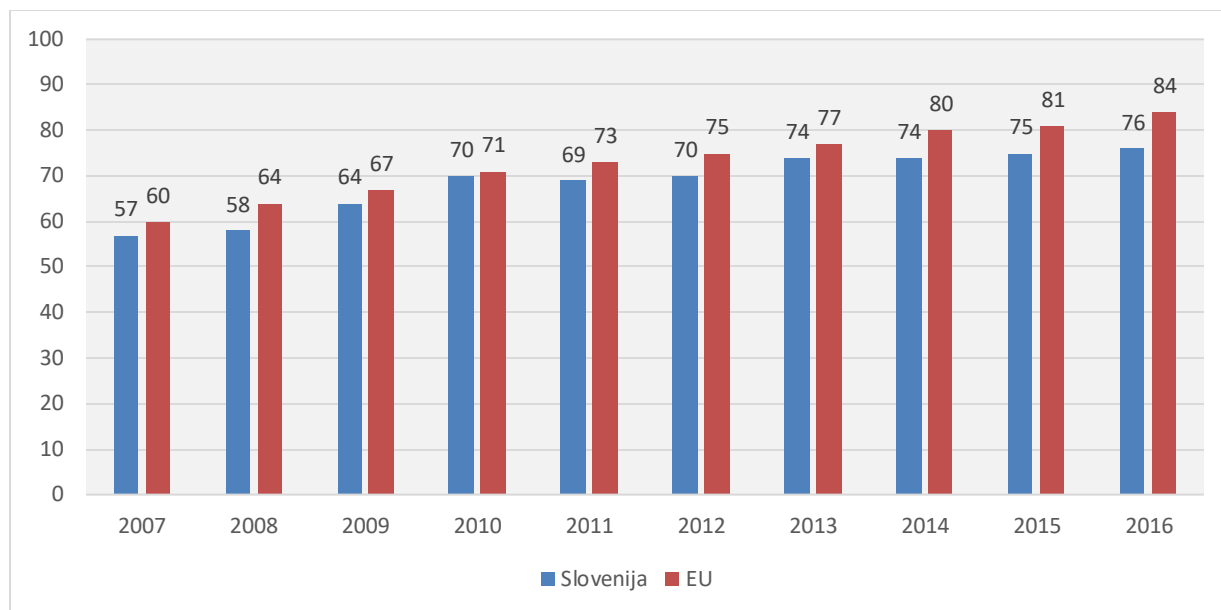
Na prelomu tisočletja je bila stopnja penetracije interneta ena od glavnih skrbi pri uporabi spletnih anket, saj je bilo število članov splošne populacije, ki jih je bilo moč doseči s spletnimi anketami, dokaj nizko (Couper 2000, 467). Danes je to število precej večje, saj so se uresničile optimistične napovedi o rasti te nove tehnologije. Po podatkih Eurostata (gl. Sliko 3.1) je imelo leta 2016 v Sloveniji dostop do interneta 78 % gospodinjstev (porast 20 odstotnih točk od leta 2007), v EU osemindvajseterici pa 85 % (porast 30 odstotnih točk od leta 2007). Ob tem je treba upoštevati razlike med posameznimi državami. Najnižjo stopnjo dostopa do interneta v EU je imela v letu 2016 Bolgarija, kjer je imelo internetno povezavo 64 % gospodinjstev, najvišjo pa Luksemburg, Nizozemska in Norveška, kjer je imelo internetno povezavo 97 % gospodinjstev. Posameznikov, ki so v zadnjih 12 mesecih dostopali do interneta, je bilo v Sloveniji 76 %, v EU pa 84 % (gl. Sliko 3.2)

Slika 3.1: Odstotek gospodinjstev z dostopom do interneta



Vir: Eurostat 2017a

Slika 3.2: Odstotek posameznikov, ki so v zadnjih 12 mesecih uporabljali internet



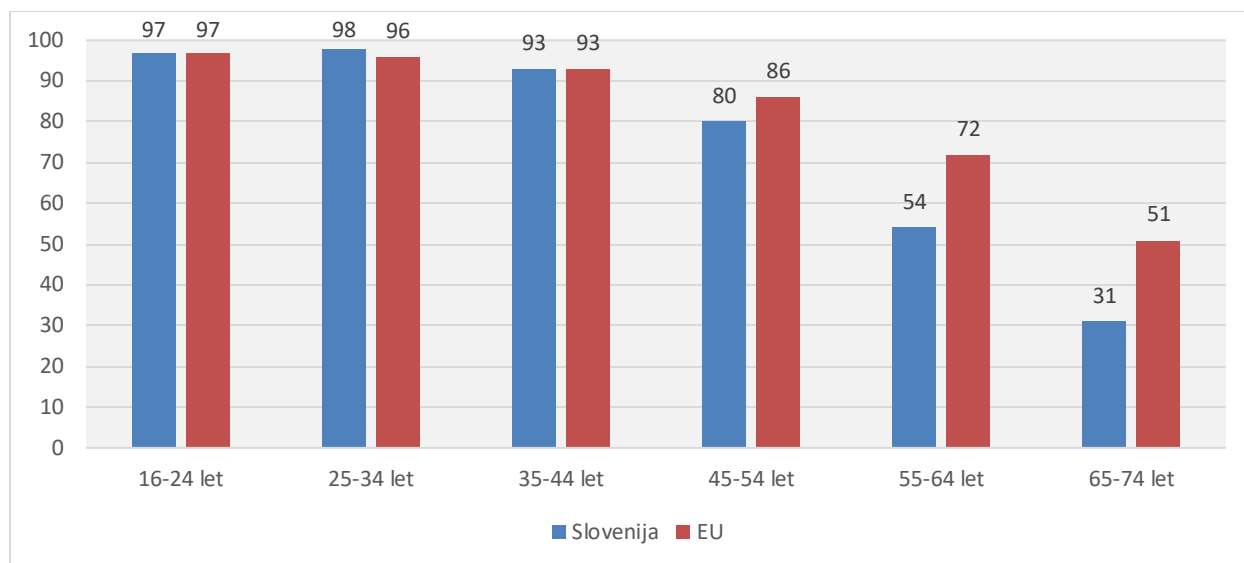
Vir: Eurostat 2017b

3.2 Razlike med posamezniki na nasprotnih straneh digitalnega razkoraka

Na prelomu tisočletja je veljalo, da so uporabniki interneta v povprečju bolj premožni, višje izobraženi in mlajši (Couper 2000, 471), podobne razlike pa se ohranjajo še danes, čeprav odstotek rednih uporabnikov interneta narašča. Posledično lastnosti uporabnikov interneta ne odražajo nujno stanja v splošni populaciji (Mohorko, de Leeuw in Hox 2011, 2). Oglejmo si Eurostatove podatke o penetraciji interneta glede na starost, spol in izobrazbo v letu 2016. Kot na Sliki 3.2 penetracija interneta v tem primeru pomeni odstotek posameznikov, ki so v zadnjih 12 mesecih uporabljali internet.

Odstotek uporabnikov interneta glede na starost v Sloveniji in EU je prikazan na Sliki 3.3, kjer je dobro viden digitalni razkorak med mlajšimi in starejšimi. Opazimo lahko, da je odstotek rabe interneta zelo visok za starostne skupine 16 do 24 let, 25 do 34 let in 35 do 44 let, nato pa začne upadati. V starostni skupini med 65 in 74 letom je uporabnikov interneta v EU samo še 51 %, v Sloveniji pa samo še 31 % (za starejše od 74 let podatkov ni bilo na voljo). Iz grafa je razvidno tudi, da je razkorak med mlajšimi in starejšimi glede na uporabo interneta v Sloveniji še precej večji, kot velja za evropsko povprečje.

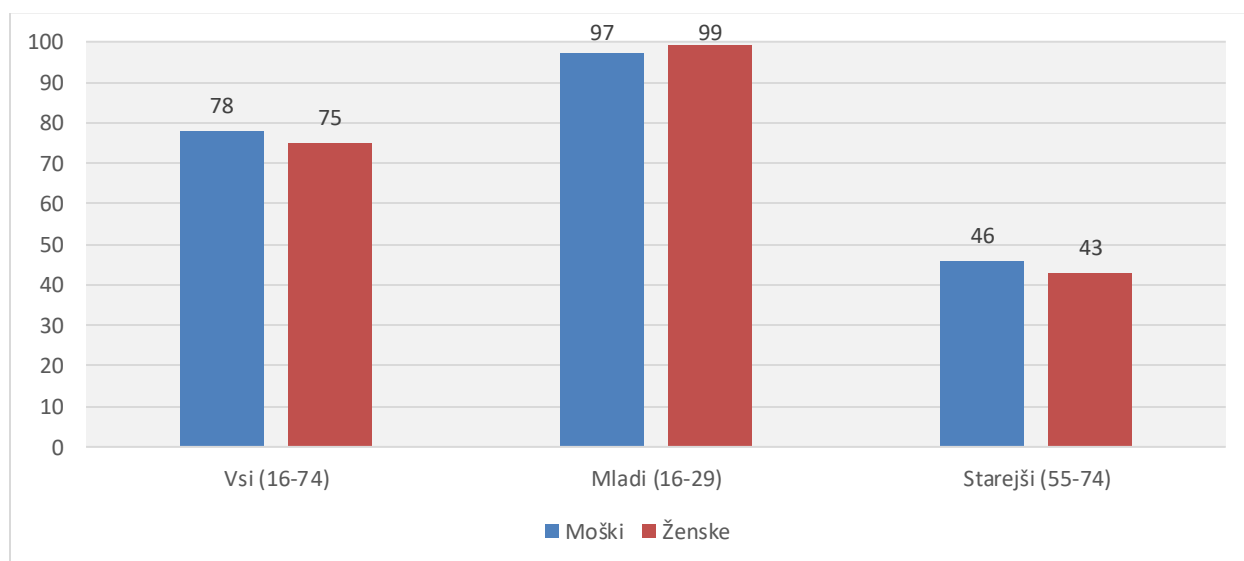
Slika 3.3: Odstotek posameznikov, ki so v zadnjih 12 mesecih uporabljali internet, po starostnih skupinah za leto 2016



Vir: Eurostat 2017b

Iz Slike 3.4 lahko razberemo, da je digitalni razkorak glede na spol v Sloveniji zelo majhen; razlika v penetraciji interneta med moškimi in ženskami je le 3 odstotne točke za celotno populacijo. Pri mlajših posameznikih med 16 in 19 letom ima dostop do interneta 97 % moških in 99 % žensk, medtem ko ima pri starejših med 55 in 74 letov dostop več moških; 46 % v primerjavi s 43 % žensk.

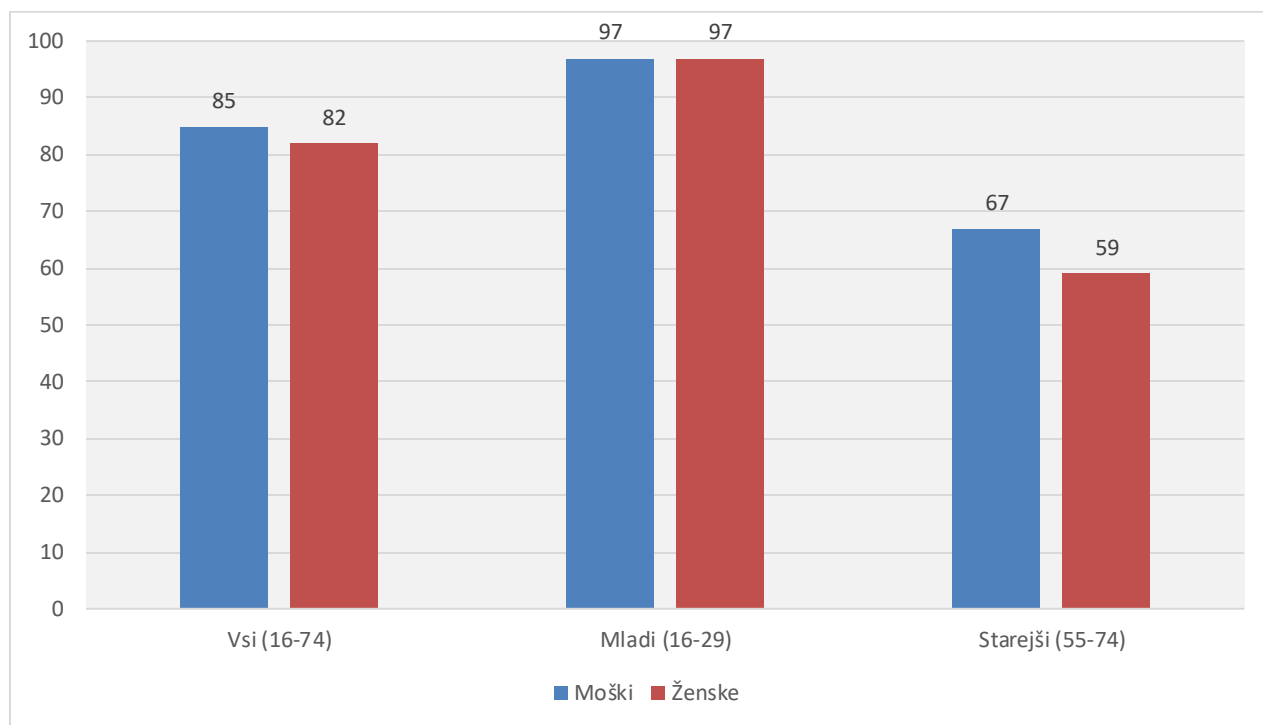
Slika 3.4: Odstotek posameznikov v Sloveniji, ki so v zadnjih 12 mesecih uporabljali internet, glede na spol in starostno skupino za leto 2016



Vir: Eurostat 2017b

Tudi v celotni EU je razlika v penetraciji 3 odstotne točke v prid moškim, vendar tu lahko opazimo večji razkorak med moškimi in ženskami pri starejših. Kot vidimo na Sliki 3.5, je v EU internetnih uporabnikov 67 % moških med 55 in 74 letom in samo 59 % žensk.

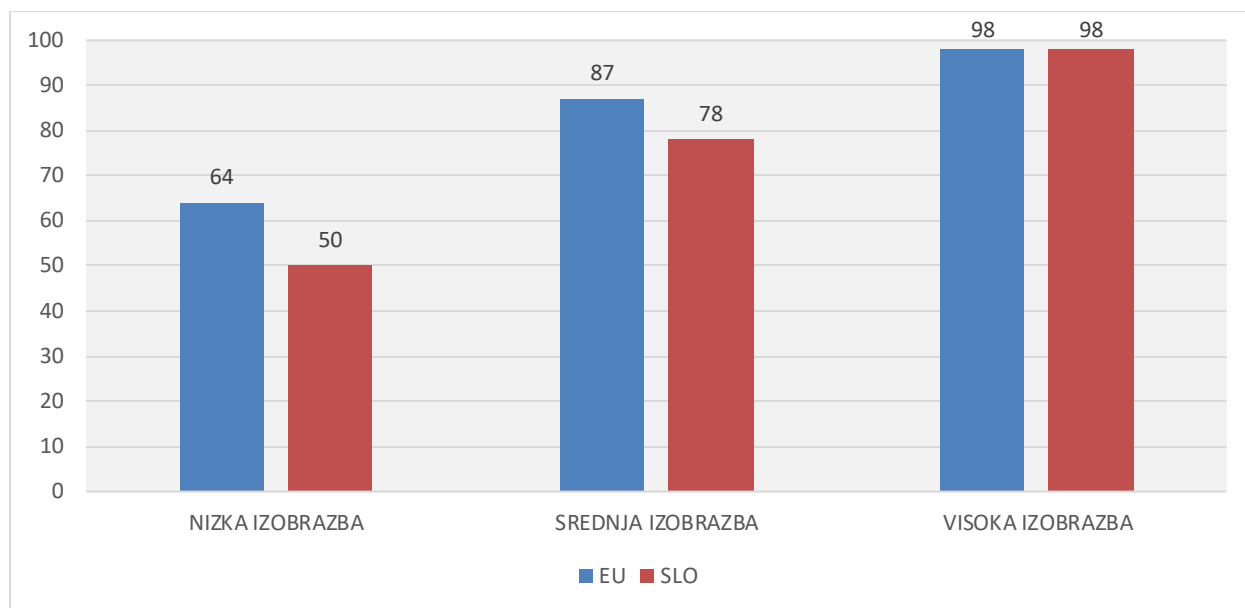
Slika 3.5: Odstotek posameznikov v EU, ki so v zadnjih 12 mesecih uporabljali internet, glede na spol in starostno skupino za leto 2016



Vir: Eurostat 2017b

Oglejmo si še razlike v penetraciji interneta glede na izobrazbo: prva skupina »nizka izobrazba« zajema posameznike z nizko formalno izobrazbo ali brez nje, druga skupina zajema posameznike s srednjo formalno izobrazbo, tretja skupina pa posameznike z visoko formalno izobrazbo. Kot lahko vidimo na Sliki 3.6, je izobrazba pomemben faktor pri digitalnem razkoraku. Tako so uporabniki interneta skoraj vsi z visoko formalno izobrazbo (98 % v Sloveniji in EU), medtem ko ta odstotek upade pri tistih s srednjo in nižjo izobrazbo. Tako ima v Sloveniji dostop do interneta samo 50 % posameznikov z nižjo izobrazbo, v EU pa 64 %.

Slika 3.6: Odstotek posameznikov v Sloveniji in EU, ki so v zadnjih 12 mesecih uporabljali internet, glede na izobrazbo



Vir: Eurostat 2017b

Digitalni razkorak je pomemben vir pristranskosti pri spletnih anketah zaradi zgoraj opisanih razlik med državami in različnimi demografskimi skupinami. Poleg razlik v demografski strukturi med uporabniki interneta in celotno populacijo pa nas skrbijo tudi potencialne razlike pri ključnih spremenljivkah, ki jih želimo meriti s spletnimi anketami. Za merjenje teh razlik je običajno potrebno izvesti paralelne ankete z različnimi metodami. Primer je raziskava Pew Research Center (Flemming and Sonner 1999 v Couper 2000), kjer so primerjali odgovore telefonske ankete z vzorcem naključno generiranih telefonskih števil, spletne ankete prostovoljnih udeležencev in spletne ankete s predhodno izbranim vzorcem anketirancev in našli pomembne razlike med temi tremi skupinami, ki se niso ujemale z nobenim predvidljivim vzorcem (Couper 2000, 471–472). Metodologija in rezultati takšnih raziskav so predstavljeni v petem poglavju.

4 Teorija anketnih napak in napake pokritja

V prejšnjih poglavjih smo spoznali značilnosti spletnih anket in digitalnega razkoraka, ki lahko pripeljejo do napak v anketnih ocenah. V nadaljevanju se posvečamo teoriji anketnih napak: kako jih definiramo in kako jih merimo. Najprej si ogledamo splošno opredelitev anketnih napak, nato pa se osredotočimo na napako pokritja in predstavimo formule za izračun pristranskosti zaradi nepokritja.

4.1 Opredelitev pojma anketne napake

Skupna anketna napaka (angl. *total survey error*) je mera, ki se nanaša na posamezno statistično karakteristiko (npr. povprečje ali delež) in ne na anketo kot celoto. Definiramo jo kot razliko

med oceno populacijskega parametra, pridobljeno z anketiranjem, in resnično vrednostjo tega parametra. Sestavljena je iz dveh komponent: vzorčna napaka (angl. *sampling error*) in nevzorčna napaka (angl. *nonsampling error*). Do vzorčne napake pride, ker v vzorec ni zajeta celotna populacija, in jo lahko zmanjšamo tako, da povečujemo vzorec. Nevzorčna napaka zajema vse vrste napak, do katerih pride med zbiranjem in obdelavo podatkov ter ocenjevanjem parametrov. Biemer in Lyberg (2003, 35–38) opisujeta pet virov nevzorčne napake. To so napaka v specifikaciji (angl. *specification error*), napaka okvira (angl. *frame error*) oz. napaka pokritja (angl. *coverage error*), napaka zaradi neodgovora (angl. *nonresponse error*), napaka merjenja (angl. *measurement error*) in napaka pri obdelavi podatkov (angl. *processing error*),

Najpogostejše uporabljeno merilo za skupno anketno napako je srednja kvadratna napaka (angl. *mean squared error* ali *MSE*). Srednja kvadratna napaka meri vpliv skupne anketne napake na izbrano oceno. Nizka srednja kvadratna napaka kaže na nizko skupno anketno napako, medtem ko visoka srednja kvadratna napaka opozarja, da ena ali več od zgoraj naštetih napak negativno vpliva na natančnost ocene. Glavni cilj anketnega načrta (angl. *survey design*) je zmanjšati srednjo kvadratno napako ob upoštevanju finančnih in časovnih omejitev specifične raziskave (Biemer in Lyberg 2003, 45).

Srednjo kvadratno napako izračunamo tako, da seštejemo kvadrate vseh komponent pristranskosti in variance, ki izvirajo iz različnih vrst napak v anketnem načrtu (Biemer in Lyberg 2003, 62). Pri tem je varianca tisti del napake, ki izvira iz slučajnih napak (angl.

variable errors), medtem ko pristranskost izvira iz sistematičnih napak (angl. *systematic errors*).

Za slučajne napake velja, da imajo kumulativni učinek na anketne odgovore. Vrednosti za specifično enoto so bodisi višje bodisi nižje, kot so v resnici, torej so napake pozitivne ali negativne. Če gledamo na podatke kot celoto, so slučajne napake pri nekaterih enotah pozitivne, pri drugih pa negativne, kar pomeni, da se do neke mere med seboj izničijo. Če so pozitivne in negativne napake enako verjetne, je skupna slučajna napaka dokaj nizka. Slučajne napake so tesno povezane s konceptom zanesljivosti podatkov (angl. *data reliability*), ki označuje razmerje med resnično variabilnostjo med člani populacije in skupno variabilnostjo v podatkih, ki vključuje tudi variabilnost zaradi slučajnih napak. Ko je razmerje zanesljivosti 1, ni slučajnih napak in vsa variabilnost izvira iz resničnih razlik med enotami. Več ko je slučajnih napak, manjše je razmerje zanesljivosti. Do sistematičnih napak pride, kadar so pozitivne napake veliko bolj pogoste od negativnih ali obratno, kar privede do napake v povprečju. To imenujemo pozitivna ali negativna pristranskost. Če povzamemo, so slučajne napake tiste, ki se med seboj izničijo, ko računamo povprečje za vsa opazovanja, sistematične napake pa tiste, ki pripeljejo do pozitivne ali negativne pristranskosti pri povprečju. Vsi viri nevzorčne napake lahko povzročajo slučajne napake, sistematične napake ali pa obe vrsti napak (Biemer in Lyberg 2003, 46–47).

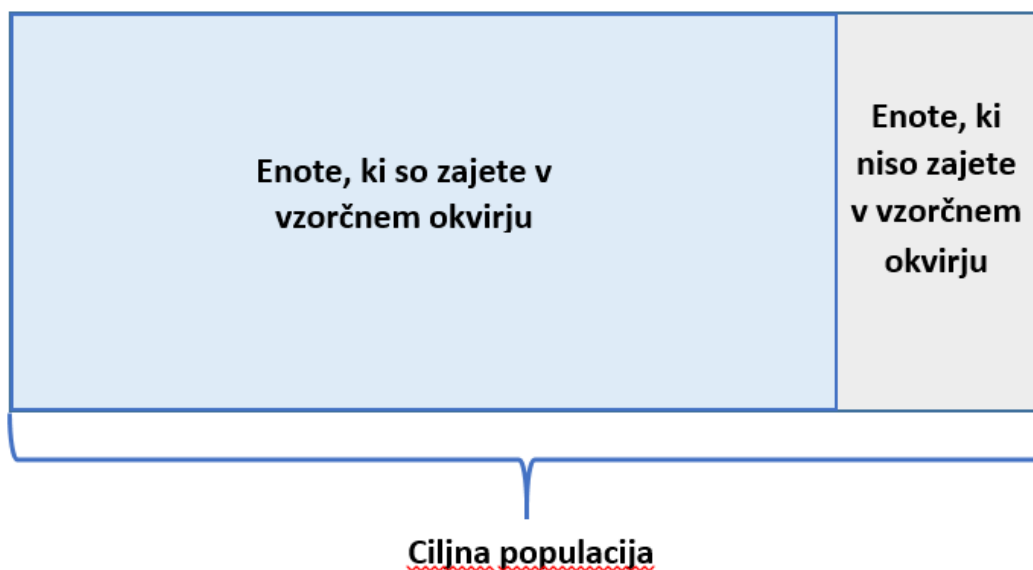
4.2 Napaka pokritja

Do napake pokritja pride v fazi načrtovanja vzorca. Ustreznost določenega vzorčnega načrta je močno odvisna od ciljev raziskave (Callegaro in drugi 2015, 41). Vzorčni načrt ima sicer več aspektov: vzorčni okvir, metoda vzorčenja, velikost vzorca in izbiranje enot za v vzorec. Idealni vzorci imajo takšno sestavo kot populacija, ki jo preučujemo, samo manjši obseg. Takšne vzorce imenujemo reprezentativni vzorci. Iz njih lahko projiciramo na populacijo tiste lastnosti, ki nas zanimajo.

Do napake pokritja pride, ko v vzorčni okvir ne uspemo vključiti vseh enot ciljne populacije. Vzorčni okvir (angl. *sample frame*) lahko predstavljajo sezname enot s kontaktnimi informacijami, zemljevidi (v primeru prostorskega vzorčenja) ali celo naključno generirane možne telefonske številke (Groves in drugi 2009, 81–83). Osnovni problem pokritja je, kako dobro ta okvir pokriva ciljno populacijo. Takoj, ko so med njima razlike, lahko pride do težav. Razlike med okvirom in populacijo so lahko različnih vrst: včasih nekatere enote iz

ciljne populacije niso zajete v okvir, včasih so v okvir zajete tudi enote, ki ne spadajo v ciljno populacijo, včasih se ista enota iz populacije v okviru pojavi več kot enkrat, včasih pa ena enota iz okvira zajema več populacijskih enot, npr. če proučujemo posamezne družinske člane, naš okvir pa je seznam gospodinjstev (Biemer in Lyberg 2003, 64–66). Biemer in Lyberg (2003, 67) poudarita, da moramo razlikovati med napakami pokritja in primeri, kjer namenoma in eksplicitno izključimo nekatere dele ciljne populacije, ki smo jo prvotno definirali, npr. zato, ker so enote težko dostopne ali pa bi dostop povzročil veliko stroškov.

Slika 4.1: Shema problema nepokritja



Vir: prirejeno po Biemer in Lyberg 2003, 64

Problem nepokritja je prikazan na Sliki 4.1. Do težave pride, ker imajo enote, ki jih ni v okviru, ničelno verjetnost, da bi bile vključene v vzorec. Delež ciljne populacije, ki ni vključen v okviru, imenujemo tudi stopnja nepokritja (angl. *noncoverage rate*). Ponavadi se stopnja nepokritja razlikuje med posameznimi regijami in populacijskimi skupinami (Biemer in Lyberg 2003, 65), kar privede do pristranskosti končnih rezultatov. Da bi presodili učinek pristranskosti zaradi nepokritja na srednjo kvadratno napako, potrebujemo dva podatka: stopnjo nepokritja in razliko med povprečjema enot, ki so vključene v okviru, in tistih, ki niso.

Kot smo že omenili, napaka pokritja spada med napake neopazovanja. Drugi tip napake neopazovanja je napaka neodgovora, do katere pride, ko za nekatere enote v vzorcu ne

uspemo dobiti odgovorov na celotno anketo ali pa na posamezna vprašanja. Obe vrsti napak neopazovanja imata enak učinek na srednjo kvadratno napako, saj v obeh primerih pride do manjkajočih podatkov, vendar so metode za zmanjševanje napak precej različne (Biemer in Lyberg 2003, 63).

S statistično formulo pristranskost zaradi nepokritja izrazimo takole (Biemer in Lyberg 2003, 69):

$$B_{nc} = (1 - t_c)(\bar{Y}_c - \bar{Y}_{nc})$$

Pristranskost zaradi nepokritja = (stopnja nepokritja) × (razlika med aritmetičnima sredinama pokrite in nepokrite populacije)

Formula za pristranskost temelji na naslednji formuli za populacijsko aritmetično sredino:

$$\bar{Y}_p = t_c \bar{Y}_c + (1 - t_c) \bar{Y}_{nc}$$

Populacijska aritmetična sredina = (stopnja pokritosti) × (aritmetična sredina pokrite populacije) + (1-stopnja pokritosti) × (aritmetična sredina nepokrite populacije)

Pristranskost zaradi nepokritja je torej funkcija stopnje nepokritja in razlike v lastnostih pokrite in nepokrite populacije (Biemer in Lyberg 2003, 77). Težava pri praktični uporabi te formule je, da nimamo podatkov o aritmetični sredini za nepokrito populacijo, razen če jih pridobimo iz nekega drugega vira ali raziskave.

Relativna pristranskost je izražena v odstotkih in jo dobimo tako, da zgoraj izračunano pristranskost delimo s populacijsko aritmetično sredino (Biemer in Lyberg 2003, 69–70):

$$RB_{nc} = (1 - t_c) \frac{\bar{Y}_c - \bar{Y}_{nc}}{\bar{Y}_p}$$

Relativna pristranskost zaradi nepokritosti = (stopnja nepokritosti) × ((razlika med aritmetičnima sredinama pokrite in nepokrite populacije) ÷ (populacijska aritmetična sredina))

Kot je mogoče razbrati iz zgornjih formul, bo v primeru zelo visoke pokritosti (tj. kadar je stopnja nepokritja blizu 0), pristranskost zaradi nepokritja zelo nizka, tudi če se pokrita in nepokrita populacija med seboj močno razlikujeta. Z dviganjem stopnje nepokritja se posledično veča tudi pristranskost, vendar je njeno naraščanje odvisno od tega, kako velike so razlike med obema populacijama. Če so si pokrite in nepokrite enote zelo podobne, bo pristranskost zaradi nepokritja ostala nizka (Biemer in Lyberg 2003, 70).

Kakšne vrednosti relativne pristranskosti so torej problematične? Biemer in Lyberg (2003, 71–72) ugotavljata, da je odgovor odvisen od ciljev raziskave in tega, kako nameravamo uporabiti podatke. Če na primer relativna pristranskost 2,4 % drastično spremeni ugotovitve ali odločitve, ki jih bomo sprejeli na podlagi podatkov, je tolikšna pristranskost še kako pomembna. Po drugi strani pa se raziskovalci lahko odločijo, da jih zanimajo samo razlike, ki so večje od 10 %. Kadar je stopnja pristranskosti nesprejemljiva, se raziskovalci lahko poslužijo dodatnih taktik, kot so zbiranje podatkov o manjkajočih enotah s pomočjo drugih okvirov in prek drugih tehnik anketiranja. Ker je takšen pristop precej drag, se lahko odločijo tudi za popravke končnih ocen (npr. uteževanje) na podlagi drugih virov.

Bethlehem in Biffignandi (2012, 289) se osredotočita specifično na pristranskost zaradi nepokritja pri spletnih anketah. To pristranskost definirata z naslednjo formulo:

$$B(\bar{y}_{HT}) = E(\bar{y}_{HT}) - \bar{Y} = \bar{Y}_I - \bar{Y} = \frac{N_{NI}}{N} (\bar{Y}_I - \bar{Y}_{NI})$$

Pristranskost = aritmetična sredina za uporabnike interneta – aritmetična sredina populacije = (Število neuporabnikov interneta ÷ število enot v populaciji) × (aritmetična sredina za uporabnike interneta – aritmetična sredina za neuporabnike interneta)

Kot v bolj splošni formuli pristranskosti zaradi nepokritja, je tudi pristranskost zaradi nepokritja neuporabnikov interneta v spletnih anketah odvisna od dveh faktorjev. Prvi je relativna velikost populacije brez dostopa do interneta, drugi pa razlike med uporabniki in neuporabniki interneta (t.i. kontrast). Pričakovati gre, da bo pristranskost zaradi prvega faktorja postala čedalje manjša z naraščanjem penetracije interneta. Po drugi strani ne moremo predvideti, kako se bo sčasoma spremenil drugi faktor. Možno je, da bodo razlike med uporabniki in neuporabniki interneta celo narasle, ko bo slednjih čedalje manj.

Bethlehem in Biffignandi (2012, 289) opozarjata, da ni nujno, da bo manjši delež neuporabnikov v kombinaciji z večjim kontrastom pripeljal do manjše pristranskosti.

Bethlehem in Biffignandi (2012, 290–292) naštejeta najpogostejše metode za zmanjševanje pristranskosti zaradi nepokritja pri spletnih anketah, ki smo jih v veliki meri obdelali že v poglavju o spletnih anketah. Prva možnost je omogočiti dostop do interneta tistim članom vzorca, ki ga še nimajo, za kar se pogosto odločijo pri longitudinalnih spletnih panelih. Druga možnost je kombiniran anketni načrt, pri katerem se rezultate spletne ankete dopolni še z drugo metodo anketiranja, npr. preko telefona. Tretji pristop je prilagoditev rezultatov spletne ankete s pomočjo uteževanja, npr. preko poststratifikacije. Pri poststratifikaciji v anketi merimo nekaj spremenljivk, za katere so znani podatki za celotno populacijo (pogosto so to spol, starost, zakonski stan in regija). Nato podatke iz ankete utežimo tako, da struktura vzorca glede na izbrane spremenljivke ustreza njihovi strukturi v ciljni populaciji.

5 Merjenje pristranskosti zaradi nepokritja v spletnih anketah

Teoretični del zaključujemo s pregledom nekaterih raziskav, ki so se ukvarjale z merjenjem pristranskosti zaradi nepokritja v spletnih anketah. Iz raziskovalnih načrtov teh raziskav lahko izluščimo metode merjenja pristranskosti zaradi nepokritja, ob enem pa nam pregled rezultatov pomaga oceniti trende glede te pristranskosti.

Formule za ocenjevanje pristranskosti zaradi nepokritja v določeni anketi, ki smo jih predstavili v prejšnjem poglavju, zahtevajo tudi dejanske vrednosti populacijskih parametrov. Natančnih vrednosti v praksi seveda nimamo, zato uporabimo podatke, pri katerih ne pričakujemo pristranskosti zaradi nepokritja. Možnosti pri tem je več, raziskave, ki jih opisujemo v tem poglavju pa vrednosti za pokrito in nepokrito populacijo (tj. uporabnike in neuporabnike interneta) črpajo iz istega niza podatkov. Ti ne izvirajo iz spletnih anket, temveč iz verjetnostnih anket splošne populacije ali pa iz verjetnostnih spletnih panelov, v katere so bili načrtno vključeni tudi neuporabniki interneta. Pri tem pristopu je ključno, da ankete vsebujejo spremenljivko, ki raziskovalcem omogoča ločevanje med uporabniki in neuporabniki interneta. Izvedejo torej simulacijo, kakšna bi bila hipotetična pristranskost, če bi izključili vse anketirance, ki niso uporabniki interneta. Te vrste raziskave tu opisujemo zato, ker tak pristop uporabimo tudi v empiričnem delu tega magistrskega dela.

5.1 Primer panela Health and Retirement Study v ZDA

Couper in drugi (2007, 132–133) so izvedli raziskavo o problemih nepokritja in neodgovora v panelni raziskavi Health and Retirement Study³ (v nadaljevanju HRS), ki pokriva prebivalce ZDA nad petdesetim letom in njihove zakonske partnerje. Njihov cilj je bilo ugotoviti, ali bi bilo možno obstoječe načine zbiranja podatkov za HRS (terenske in telefonske intervjuje) dopolniti tudi s spletnimi anketami.

Za začetek so identificirali udeležence panelne študije HRS, ki redno uporabljajo internet. Uporaba interneta je bilo eno od vprašanj v valu anketiranja leta 2002. Tisti, ki so se opredelili kot redni uporabniki, so dobili še dodatno vprašanje, ali bi bili pripravljeni sodelovati v anketi preko interneta. Vzorec za raziskavo je bil zajet iz tistih uporabnikov, ki so rekli, da bi bili pripravljeni sodelovati. Tem anketirancem so po pošti poslali vabila, ki so vsebovala povezavo do ankete in podatke za prijavo. Tistim, ki niso odgovorili, so poslali do

³ Ime Health and Retirement Survey bi lahko prevedli kot Raziskava zdravje in upokožitev.

tri pisma z opomniki, preko pisem pa so kontaktirali tudi tiste, ki so na spletno anketo začeli odgovarjati, a je niso končali.

Od 16077 anketirancev, ki so odgovorili na vprašanje o uporabi interneta iz leta 2002, je bilo rednih uporabnikov 29,7 %, kar se je ujemalo z drugimi podatki o penetraciji interneta pri tej starostni skupini. Raziskovalce so poleg demografskih razlik zanimale tudi razlike pri spremenljivkah, povezanih z zdravjem in ekonomskim statusom uporabnikov in neuporabnikov interneta.

Demografske razlike so preučevali z modelom, ki je uporabil naslednje spremenljivke: regija, urbanost okolja, starost, rasa oz. etnična pripadnost, spol, izobrazba, zakonski stan in zaposlenost. Rezultati so pokazali, da je uporaba interneta statistično značilno višja na urbanih območjih, med višje izobraženimi, med belci, ki niso latinskega porekla, med poročenimi, zaposlenimi in prebivalci regije Zahod. Razlik med spoloma niso našli.

Ko so v model vključili še vprašanja o financah in zdravju, so ugotovili, da so te spremenljivke statistično značilno povezane z uporabo interneta tudi, če se kontrolira za vpliv demografskih spremenljivk. Ta ugotovitev se ne sklada z argumentom, da bi izločanje vpliva demografskih spremenljivk odpravilo večino pristranskosti, ki je rezultat razlik v dostopu do interneta. Ugotovili so, da je raba interneta pozitivno korelirana z dohodkom gospodinjstva, bivanjem v enodružinski hiši (namesto v mobilnih domovih ali stanovanjih) in višjo samooceno zdravja, spomina in funkcioniranja. Uporabniki interneta nad 65 letom veliko bolj verjetno še vedno lahko vozijo in poročajo o boljšem vidu. Poleg tega so manj verjetno trenutno kadilci, v večjem številu poročajo, da so kdaj pili alkohol, po stopnji telesne aktivnosti pa se ne razlikujejo od neuporabnikov interneta.

Couper in drugi (2007, 139) zaključujejo, da vzorec, ki bi zajel samo uporabnike interneta v populaciji, ki je vključena v HRS, ne bi bil reprezentativen, niti če bi kontrolirali za demografske razlike v rabi interneta, saj so razlike med uporabniki in neuporabniki precejšnje. Ob enem predvidevajo, da bodo te razlike verjetno upadle z leti, ko se bo kohortu HRS pridružilo čedalje več oseb, ki so odrasle z internetom. Dopuščajo tudi možnost uporabe interneta kot dodatne metode zbiranja informacij v velikih panelnih študijah, kot je HRS, saj imajo te študije na voljo veliko informacij, s katerimi je mogoče v določenih okoliščinah rezultate naknadno prilagoditi.

5.2 Primer mobilnih spletnih anket v Evropi

Problem nepokritja je zelo aktualen tudi pri mobilnih spletnih anketah (angl. *mobile web surveys*). To so spletne samoankete, ki se izpolnjujejo preko pametnih telefonov. Ta metoda ima veliko skupnega s spletnimi anketami na splošno, prinaša pa nekaj specifik, kot je npr. možnost vzorčenja z naključnim generiranjem mobilnih telefonskih števil (Fuchs in Busse 2009, 22)

Fuchs in Busse sta leta 2009 izvedla raziskavo o pristranskosti zaradi nepokritja v mobilnih spletnih anketah. Od takrat sta se penetracija pametnih telefonov in uporaba mobilnega interneta močno povečali, tako da nas bodo pri tej raziskavi zanimali predvsem metodološki pristop k ocenjevanju pristranskosti zaradi nepokritja in splošne ugotovitve.

Fuchs in Busse (2009, 23) sta uporabila podatke iz Evrobarometra, kar bomo storili tudi v naši empirični analizi. Evrobarometer zajema prebivalce držav EU in držav kandidatk za članstvo v EU. V okviru te študije je vsako leto izvedenih od 4 do 8 valov terenskih anket, verjetnostno vzorčenje pa se izvede za vsako državo posebej. Evrobarometrov vprašalnik vključuje sociodemografski modul, ki je enak za vsak val ankete, module, ki se med seboj izmenjujejo v različnih valovih, in pa vprašanja, ki se pojavijo samo v enem valu. Fuchs in Busse sta uporabila podatke iz let 2005 in 2007, saj so vključevali vprašanje o dostopu do mobilnega spleta. S pomočjo tega vprašanja in vprašanja o dostopu do klasičnega interneta sta se lahko osredotočila na tri skupine anketirancev: populacijo z mobilnim internetom, populacijo s stacionarnim internetom in splošno populacijo.

Pristranskost zaradi nepokritja sta ocenjevala na podlagi spremenljivk, za katere sta sklepala, da so povezane z vsebinsko zanimivimi spremenljivkami v akademskem in tržnem raziskovanju. To so bile spremenljivke zakonski stan, starost, starost pri najvišji doseženi stopnji izobrazbe, spol in tip skupnosti. Za vsako od teh spremenljivk sta izbrala eno vrednost (npr. poročene anketirance pri spremenljivki zakonski stan) in ocenila relativno pristranskost zaradi nepokritja za to skupino, pri čemer sta uporabila formulo za relativno pristranskost, ki jo opisujeta Biemer & Lyberg (2003) in smo jo predstavili v poglavju 4.1. Pri takšnem pristopu gre za simulacijo, saj se ocenjuje potencialna pristranskost zaradi nepokritja pri hipotetičnih mobilnih spletnih anketah, medtem ko sami podatki prihajajo iz terenskih anket, na katere ne vpliva pristranskost zaradi nepokritja. Čeprav pristranskost zaradi nepokritja pri

Evrobarometru ni problematična, je vseeno treba upoštevati potencialno pristranskost zaradi vzorčne napake in neodgovorov (Fuchs in Busse 2009, 24).

Kljub dokaj visoki pristranskosti zaradi nepokritja pri sociodemografskih spremenljivkah za uporabnike mobilnega interneta sta Fuchs in Busse (2009, 21) ugotovila, da je bila ponekod po Evropi ta pristranskost že nižja kot pristranskost zaradi nepokritja pri uporabnikih stacionarnega interneta. Relativna pristranskost zaradi nepokritja pri uporabnikih mobilnega interneta se je močno razlikovala med različnimi državami. Od sociodemografskih spremenljivk je bilo največ pristranskosti pri starosti, sledila pa ji je izobrazba, medtem ko je bila pristranskost za zakonski stan in spol precej nižja.

5.3 Primer panela LISS

Nizozemska študija LISS (»Longitudinal Internet Studies for the Social Sciences«) se izvaja od leta 2007 in je med bolj znanimi verjetnostnimi spletnimi paneli v Evropi. Zajema 4500 gospodinjstev (skupno 7000 posameznikov), ki so bila naključno zajeta iz nizozemske populacije. Da bi preprečili pristranskost zaradi nepokritja, so v okviru te študije gospodinjstvom brez dostopa do interneta priskrbeli širokopasovno internetno povezavo in enostaven osebni računalnik SimPC, preko katerega lahko izpolnjujejo spletne ankete. Ta naprava je preprosta za uporabo in prilagojena tudi za starejše uporabnike (LISS Panel Data 2009). Velik začetni vložek v vključevanje gospodinjstev brez interneta pomeni, da so podatki zbrani v okviru študije LISS zelo uporabni za preučevanje pristranskosti zaradi nepokritja, saj raziskovalci lahko primerjajo uporabnike in neuporabnike interneta in vidijo, kaj bi se zgodilo, če bi slednje izločili. V nadaljevanju si bomo ogledali metodologijo in rezultate treh raziskav nepokritja na podlagi podatkov študije LISS.

5.3.1 Leenheer in Scherpenzeel (2013)

Leenheer in Scherpenzeel (2013, 18) se ukvarjata z raziskovalnim vprašanjem, ali se kakovost spletnega panela pomembno poveča, če priskrbimo dostop do interneta gospodinjstvom, ki ga nimajo. Zanima ju tudi dolgoročna kakovost panela, če se odločimo za to investicijo.

Prvi korak raziskave je ugotavljanje razlik med gospodinjstvi, ki imajo dostop do interneta, in tistimi, ki dostopa nimajo. Leenheer in Scherpenzeel (2013, 20–21) uporabita logistični regresijski model, v katerem je dostop do interneta pojasnjen s sociodemografskimi

značilnostmi gospodinjstev. Glavne ugotovitve pokažejo, da dostopa do interneta pogosto ni v gospodinjstvih, ki vključujejo starejše, imigrante, ki ne prihajajo iz zahodnih držav, oziroma samo enega člana, ali pa se nahajajo v regijah, ki imajo visok ali pa nizek nivo urbanizacije. Takšna gospodinjstva posledično manj pogosto sodelujejo v spletnih panelih.

V drugem koraku se Leenheer in Scherpenzeel (2013, 21) osredotočita na reprezentativnost vzorca gospodinjstev brez interneta, ki so bila rekrutirana v panel LISS. To ocenita s pomočjo logističnega regresijskega modela za sodelovanje gospodinjstev brez interneta v vzorčnem okviru za panel LISS. Koeficienti, ki niso statistično značilni, nakazujejo reprezentativnost za skupino neuporabnikov interneta, ki jih definira specifična pojasnjevalna spremenljivka. Takšni so koeficienti pri spremenljivkah velikost, izvor in nivo urbanizacije gospodinjstva, kar nakazuje, da te ne vplivajo na verjetnost sodelovanja v panelu. Statistično značilen učinek ima spremenljivka starost, kjer verjetnost sodelovanja narašča do povprečne starosti 31,1, nato pa spet začne upadati. Če povzamemo, vzorec gospodinjstev brez interneta v panelu LISS je reprezentativen glede na etnični izvor, regijo in število članov gospodinjstva (ena oseba ali več oseb). Vzorec ni popolnoma reprezentativen glede na starost, saj so starejše skupine premalo zastopane.

V tretjem koraku Leenheer in Scherpenzeel (2013, 22–23) ocenjujeta zvestobo panelu pri neuporabnikih interneta. Ta je višja kot pri uporabnikih interneta (torej tistih, ki niso potrebovali naprave SimPC), saj so neuporabniki izpolnili večje število ponujenih vprašalnikov (69,8 % v primerjavi z 59,8 % pri uporabnikih) in imeli nižjo stopnjo osipa (angl. *attrition rate*). Med aprilom 2008 in januarjem 2010 je panel zapustilo 9,9 % gospodinjstev brez interneta in 13,8 % gospodinjstev z internetom. Z logističnim modelom preverjata, ali je osip višji pri specifičnih podskupinah. Tudi ta model kaže, da je osip bolj verjeten za gospodinjstva z internetom. Manj verjeten je za gospodinjstva s člani, ki ne prihajajo iz zahodnih držav, bolj verjeten pa za gospodinjstva z višjo povprečno starostjo. Če se omejimo samo na gospodinjstva brez interneta, te povezave med starostjo in osipom ne najdemo. V gospodinjstvih brez interneta je osip bolj verjeten, če so enočlanska, drugih pomembnih učinkov pa ni opaziti.

Četrty korak raziskave je odgovoriti na vprašanje, ali sodelovanje v spletnem panelu vpliva na uporabo interneta pri gospodinjstvih, ki ga na začetku niso uporabljala. Naprava SimPC omogoča uporabo interneta tudi v druge namene, kar lahko sčasoma pripelje do čedalje večje

podobnosti med tistimi gospodinjstvi, ki so že od začetka imela internet, in uporabniki SimPC. Primerjava uporabe interneta med uporabniki in neuporabniki (torej uporabniki SimPC), ki so na internetu preživeli vsaj eno uro tedensko, pokaže, da člani gospodinjstev z internetom bolj pogosto uporabljajo internet kot člani gospodinjstev brez interneta za vse namene razen za branje novic in revij. V obeh skupinah večina aktivnih uporabnikov interneta uporablja e-pošto (98,2 % v gospodinjstvih z internetom in 91,5 % v gospodinjstvih brez). Velike razlike najdemo pri uporabi spletnega bančništva (83,8 % proti 52,0 %), spletnega nakupovanja (67,9 % proti 34,9 %) in prenašanja glasbe in filmov (31,1 % proti 12,2 %). Vidimo lahko, da neuporabniki, ki začnejo uporabljati internet, uporabljajo manj različnih možnosti, ki so v večini bolj preproste za uporabo (Leenheer in Scherpenzeel 2013, 25–26).

Kar se tiče vprašanja reprezentativnosti, primerjava sestave panela s celotno Nizozemsko populacijo pokaže, da je panel bolj podoben splošni populaciji, če so v njem vključena gospodinjstva brez interneta. Pri gospodinjstvih nad starostjo 65 let je odstotek v splošni populaciji 14,8 %, medtem ko je ta številka pri internetnih gospodinjstvih v panelu LISS le 9,3 %, če vključimo gospodinjstva brez interneta pa doseže 11,2 %. V celotnem panelu so bolje zastopana tudi enočlanska gospodinjstva in gospodinjstva, ki nimajo lastnega prebivališča, kar bolj ustreza sliki splošne populacije (Leenheer in Scherpenzeel 2013, 24).

Leenheer in Scherpenzeel (2013, 18) opozarjata tudi na problem nizke celotne začetne stopnje odgovora (angl. *total initial response rate*) pri panelnih študijah, kjer se neuporabnikom priskrbi dostop do tehnologije. Hoogendoorn in Daalmans (2009 v Leenheer in Scherpenzeel 2013) sta naletela na veliko nižjo stopnjo odgovora med potencialnimi anketiranci, ki niso imeli osebnega računalnika, kot med tistimi, ki so ga že imeli, čeprav zaključujeta, da je tudi ta nizka stopnja boljša, kot če jim sodelovanja sploh ne bi ponudili. Tudi Leenheer in Scherpenzeel sta odkrila podobne razlike za panel LISS. Stopnja rekrutiranja (angl. *recruitment rate*) je bila za gospodinjstva z dostopom do interneta 84 %, za gospodinjstva brez dostopa pa 35 %. Postavlja se torej vprašanje, ali z vključevanjem gospodinjstev brez interneta dovolj povečamo reprezentativnost ob upoštevanju pristranskosti zaradi začetnega neodgovora.

5.3.2 Eckman (2015)

Eckman (2015, 41) je za pet objavljenih člankov, ki temeljijo na podatkih iz študije LISS, preverila, kako bi se njihovi zaključki spremenili, če v študijo ne bi vključili gospodinjstev brez dostopa do interneta⁴. V ta namen je replicirala tako opisne statistike kot tudi regresijske modele iz člankov na vzorcih z neuporabniki interneta in brez njih. Raziskave, ki so replicirane v članku, spadajo na različna znanstvena področja, vse pa so bile objavljene v recenzirani publikaciji, uporabljajo podatke iz študije LISS in poleg vsaj enega regresijskega modela vključujejo tudi tabelo z opisno statistiko (Eckman 2015, 44).

V prvem koraku analize Eckman (2015, 45) oceni populacijske aritmetične sredine za vse spremenljivke, ki so uporabljene v modelih, najprej na celotnih podatkih, nato pa še na podatkih brez neuporabnikov interneta. Ocena pristranskosti zaradi nepokritja za vsako spremenljivko je razlika med drugo in prvo aritmetično sredino. Da bi preverila, ali je izračunana pristranskost statistično značilna, uporabi linearni model, v katerem je spremenljivka Y, ki nas zanima, odvisna, neodvisna spremenljivka pa je indikator, ali je gospodinjstvo prejelo SimPC (torej indikator, da gre za neuporabnike interneta). Da bi pristranskost lažje primerjali med spremenljivkami z različnimi enotami, izračuna še absolutno relativno pristranskost, ki je izražena v odstotkih.

V naslednjem koraku replicira regresijske modele iz člankov in jih ponovno oceni še s podatki brez neuporabnikov interneta. Nato uporabi merilo prekrivanja (angl. *overlap measure*), ki so ga razvili Karr, Kohnen, Ognian, Reiter, in Sanil (2006), da bi učinek odstranitve neuporabnikov interneta predstavila bolj kvantitativno. Merilo prekrivanja oceni s primerjavo intervalov zaupanja za koeficiente pri celotnem vzorcu in vzorcu brez neuporabnikov interneta, da bi videla, kolikšen delež intervala zaupanja za eno skupino leži na stičišču intervalov zaupanja obeh skupin. Merilo zavzema vrednosti med 0 % (brez prekrivanja intervalov zaupanja) in 100 % (intervala zaupanja se popolnoma prekrivata). Višji ko je odstotek, manj se model spremeni, če izzamemo neuporabnike interneta (Eckman 2015, 46).

⁴ Eckman za gospodinjstva brez interneta, ki so prejela SimPC, da bi lahko sodelovala v panelu, uporablja izraz podprta gospodinjstva (angl. *supported households*), vendar v tem delu ostajamo pri izrazu neuporabniki, ki se je uveljavil v prejšnjih poglavjih.

Nazadnje si ogleda še, kako z neodvisnimi spremenljivkami v logističnem regresijskem modelu lahko predvidimo, ali je enota uporabnik ali neuporabnik interneta. Uporaba kontrolnih spremenljivk, ki korelirajo z indikatorjem, ali je enota neuporabnik interneta, pomaga zmanjšati vpliv pristranskosti zaradi nepokritja na ocene koeficientov. Tudi to merilo zavzema vrednosti med 0 % in 100 %, pri čemer višje vrednosti pomenijo, da model bolje razlikuje med uporabniki in neuporabniki (Eckman 2015, 46).

Rezultati so pokazali, da se gospodinjstva brez interneta od drugih močno razlikujejo po demografskih značilnostih in da bi njihova izključitev pripeljala do pristranskosti ocene povprečij nekaterih spremenljivk. Kljub temu nepokritje neuporabnikov interneta v opisanih raziskavah povečini ne bi pripeljalo do pristranskosti v ocenah modelov, interpretacija rezultatov v repliciranih člankih pa bi najverjetneje ostala enaka (morebitna izjema je samo eden od člankov, kjer ena od dveh spremenljivk, na kateri se pri interpretaciji zanaša avtor, ni več statistično značilna, če izključimo neuporabnike interneta). Manjše tveganje pristranskosti zaradi nepokritja lahko pojasnimo z nizko stopnjo nepokritja in z vključitvijo kontrolnih spremenljivk, ki so korelirane z nepokritjem (Eckman 2015, 55).

Eckman (2015, 55) opozarja, da moramo biti previdni pri izpeljevanju širših zaključkov, saj so se replicirane analize med seboj močno razlikovale. Nekatere so se ukvarjale z odvisnimi spremenljivkami, ki so močno povezane z dostopom do interneta (npr. raziskava o internetnih goljufijah), v drugih pa odvisne spremenljivke z internetom sploh niso bile povezane (npr. raziskava o odnosih med očeti in sinovi). Zaključki bi bili lahko drugačni pri drugih člankih, ki temeljijo na podatkih iz študije LISS, prav tako pa ne izključujejo možnosti pristranskosti zaradi nepokritja pri podatkih za druge spletne ankete in druge države. Pristranskost zaradi nepokritja je bolj verjetna pri spletnih anketah v državah z nižjo penetracijo interneta kot na Nizozemskem in pa pri študijah, ki se osredotočajo na starejšo populacijo ali migrante, razen če izvajalci zagotovijo dostop do interneta tudi neuporabnikom iz teh skupin. Za izvajalce anket, ki si vključitve neuporabnikov interneta ne morejo privoščiti, raziskava podaja naslednje ugotovitve. Pristranskost zaradi nepokritja pogosteje vpliva na ocene povprečij, medtem ko so ocene iz multivariatnih modelov bolj robustne. Pred pristranskostjo zaradi nepokritja se je mogoče nekoliko zaščititi, če avtorji spletnih anket merijo tudi spremenljivke, ki so korelirane z dostopom do interneta, in jih upoštevajo v končnem modelu.

5.3.3 Toepoel in Hendriks (2016)

Toepoel in Hendriks (2016) sta razširila raziskavo Eckman (2015). Preučevala sta učinek izključitve neuporabnikov interneta iz vzorca za spletne ankete na podlagi podatkov iz študije LISS med leti 2008 in 2013. Zanimale so ju razlike med predhodnimi uporabniki interneta in uporabniki naprave SimPC v sociodemografskih značilnostih in odgovorih na vprašalnike z različno tematiko. Poleg tega sta raziskala še, kako vključitev uporabnikov SimPC vpliva na povezave med spremenljivkami.

Spremenljivke, ki so bile vključene v raziskavo, so bile spol, starost, zakonski stan, poklic, izobrazba, dohodek in poreklo. Primerjala sta tudi uporabo spleta med uporabniki SimPC in klasičnimi internetnimi uporabniki in odstotek uporabnikov, ki na internetu preživijo vsaj eno uro tedensko. S tem sta želela ugotoviti, ali odstotek uporabnikov SimPC, ki redno uporabljajo internet, z leti narašča. Poleg tega sta vključila tudi izbor spremenljivk iz osmih osrednjih modulov študije LISS: zdravje, vera in etnično poreklo, ekonomska situacija in nastanitev, politika in vrednote, osebnost in vključenost v družbo in prosti čas.

Toepoel in Hendriks (2016, 36) sta primerjala odstotek uporabnikov interneta v panelu LISS s podatki nizozemskega statističnega urada CBS o dostopu do interneta med prebivalci Nizozemske. Ugotovila sta, da se reprezentativnost panela spreminja z leti, vendar razlika ne preseže 6 odstotnih točk. Ko so rekrutirali udeležence za raziskavo LISS, so dobili manjši odziv pri težje dosegljivih skupinah, kot so starejši, kar lahko pojasni začetno razliko, kjer so bili redni uporabniki interneta vključeni v prevelikem številu (angl. *overrepresented*). V kasnejših letih so postali vključeni v premajhnem številu (angl. *underrepresented*), saj so veliko pogosteje zapustili panel kot uporabniki SimPC, medtem pa je penetracija interneta na Nizozemskem še naprej rasla.

Demografske razlike med uporabniki in neuporabniki interneta sta ocenila z uporabo logistične regresije, ki preverja, katere demografske spremenljivke vplivajo na spremenljivko Uporaba SimPC. Ena od ugotovitev je bila, da so bili člani panela mlajši od 24 let manj verjetno uporabniki SimPC od referenčne skupine v starosti 25–44 let, medtem ko je bilo za starejše anketirance to bolj verjetno. S starostjo se torej povečuje verjetnost uporabe SimPC, torej verjetnost, da gre za tiste, ki so bili pred vključitvijo v panel neuporabniki interneta.

Vpliv uporabe SimPC na rabo spleta (tj. koliko ur v povprečju uporabniki preživijo na spletu) sta ocenila z enofaktorsko analizo variance, pri kateri sta kontrolirala sociodemografske značilnosti. Redni uporabniki so internet uporabljali povprečno 17 ur na teden, uporabniki SimPC pa 9 ur. Ugotovila sta, da uporaba SimPC nima pomembnega vpliva na rabo interneta, saj so razlike pojasnjene s sociodemografskimi značilnostmi. Obenem sta ugotovila, da odstotek uporabnikov SimPC, ki na internetu preživijo vsaj eno uro tedensko (poleg sodelovanja v panelu) z leti narašča, medtem ko ostaja enak za redne uporabnike interneta.

Pri ključnih spremenljivkah iz osrednjih modulov sta našla razlike na področju zdravja, vere in etničnega porekla, ekonomske situacije in nastanitve ter politike in vrednot, medtem ko se neuporabniki in uporabniki niso razlikovali pri modulih osebnost in družbena integracija ter prosti čas.

Rezultati študije so pokazali, da se uporabniki naprave SimPC od rednih uporabnikov interneta razlikujejo pri večini sociodemografskih značilnosti in pri več raziskovanih spremenljivkah, vendar je število teh uporabnikov tako majhno, da izključitev iz študije nima pomembnega vpliva na končni rezultat. To kaže, da je vpliv pristranskosti zaradi nepokritja najbolj odvisen od tega, kako veliko je število neuporabnikov interneta v dani populaciji. Na podlagi tega odkritja avtorja zaključujeta, da se v državi s tako visoko penetracijo interneta, kot jo ima Nizozemska, ne izplača priskrbeti internetnega dostopa gospodinjstvom, ki ga še nimajo, razen če raziskovalce posebej zanima ta skupina prebivalcev. Obenem opozarjata na etične in znanstvene probleme z izključevanjem neuporabnikov interneta.

6 Empirična analiza podatkov iz Evrobarometra

Empirični del tega magistrskega dela se zgleduje po analizi podatkov iz Evrobarometra med leti 2005 in 2009, ki so jo izvedli Mohorko in drugi (2011), da bi ocenili pričakovano pristranskost zaradi nepokritja pri spletnih anketah v Evropi. Evrobarometer zbira podatke preko terenskih anket in vključuje vprašanje o dostopu do interneta, torej pokriva tako uporabnike kot tudi neuporabnike interneta in hkrati omogoča primerjavo med njimi (Mohorko in drugi 2011, 3).

Mohorko in drugi (2011, 5) so za ocenjevanje pristranskosti izbrali tri demografske spremenljivke (spol, starost in starost ob koncu izobraževanja⁵) in dve sociopolitični spremenljivki (politična usmeritev levo-desno in zadovoljstvo z življenjem). Izvedli so tudi večnivojsko analizo (angl. *multilevel analysis*), s katero so želeli prikazati in pojasniti trende glede na leto in državo za vseh pet indikatorjev pristranskosti. Poleg spremenljivk, ki se nanašajo na posameznike in so jih pridobili iz Evrobarometra, so uporabili tudi socioekonomske spremenljivke za države. Izbrane spremenljivke na nivoju držav so bile pričakovana življenjska doba, izobraževalni indeks, trajanje primarne in sekundarne izobrazbe v letih, urbanizacija, odstotek delovne sile, Ginijev koeficient ekonomske neenakosti, bruto domači proizvod in inflacija. Večnivojskega modela v tem magistrskem delu ne bomo replicirali.

Rezultati analize so pokazali, da se je penetracija interneta z leti zvišala v vseh državah, ki so bile vključene v študijo, posledično pa se je znižala tudi absolutna relativna pristranskost za štiri od petih spremenljivk (ta trend se ne pojavi pri spremenljivki politična usmeritev). Večnivojska analiza je pokazala, da se hitrost tega krčenja razlik razlikuje med državami. Spremenljivke na nivoju držav povsem pojasnijo te variacije pri spremenljivkah starost in zadovoljstvo z življenjem, le delno pa pri spremenljivkah spol in starost pri končani izobrazbi. Nižja pristranskost zaradi nepokritja je povezana z višjimi nivoji ekonomske razvitosti, izobrazbe in zdravja v določeni državi, medtem ko je višja pristranskost povezana z večjimi razlikami v dohodku. Avtorji zaključujejo, da pristranskost zaradi nepokritja

⁵ Čeprav Mohorko in drugi (2011) to spremenljivko imenujejo dolžina izobraževanja, vprašanje v Evrobarometru sprašuje, koliko so bili anketiranci stari ob koncu rednega izobraževanja. Tako oblikovano vprašanje ne omogoča razlikovanja med različnimi stopnjami izobrazbe (npr. med anketiranci, ki so pri starosti 30 let zaključili prvo stopnjo študija, in tistimi, ki so pri isti starosti zaključili doktorat), poleg tega pa kodina knjiga kot najnižjo veljavno vrednost sprejema 2 leti, kar ni smiselno glede na vprašanje. Ker so zelo nizke in zelo visoke vrednosti te spremenljivke v praksi precej redke, nimajo bistvenega vpliva na povprečje, vendar moramo biti previdni pri enačenju višjih vrednosti te spremenljivke z višjo izobraženostjo.

uporabnikov interneta v evropskih državah izginja, vendar ne povsod enako hitro (Mohorko in drugi 2011, 14).

6.1 Raziskovalno vprašanje

Glavno raziskovalno vprašanje tega magistrskega dela je, *kakšen bi bil vpliv na rezultate anket na osnovi vzorcev, če le-ti vključujejo ali ne vključujejo uporabnikov interneta*. Prvi primer takšnih vzorcev bi lahko bili verjetnostni spletni paneli, kjer neuporabniki interneta v zameno za sodelovanje dobijo opremo in dostop do interneta (takšni, kot jih v praksi imamo sedaj), drugi primer pa verjetnostni spletni paneli, kjer so neuporabniki interneta preprosto izključeni iz sodelovanja. Da bi odgovorili na to vprašanje, bomo primerjali Evrobarometrove anketirance, ki imajo doma internetno povezavo, s tistimi, ki je nimajo, in ocenili hipotetično pristranskost zaradi nepokritja, če bi bili v raziskavo vključeni samo uporabniki interneta s povezavo od doma. Ob enem nas zanima tudi, kako so se ocene pristranskosti zaradi nepokritja spremenile od obdobja med leti 2005 in 2009, ki ga pokriva analiza Mohorko in drugih (2011), do naslednjega obdobja med leti 2010 in 2016, ki ga pokrivamo v tej analizi. Glede na rast penetracije interneta, ki smo jo predstavili v teoretičnem delu, želimo ugotoviti, ali je še vedno smiselno investirati v to, da v spletne panele vključujemo udeležence brez internetnega dostopa (npr. preko posredovanja tablic in plačevanja internetnega dostopa neuporabnikom).

6.2 Podatki in uteževanje

Analiza temelji na podatkih iz Evrobarometra za leta 2010 do 2016. Evrobarometrove ankete se izvajajo terensko na večstopenjskem verjetnostnem vzorcu, zajetem iz splošne populacije, tako da naj ne bi prišlo do pristranskosti podatkov zaradi nepokritja. Vseeno se v rezultatih lahko pojavi pristranskost, npr. zaradi neodgovora. Ta se razlikuje med državami, ki sodelujejo v Evrobarometru, in je lahko precej visoka, čeprav točni podatki niso objavljeni. Ta problem deloma rešuje zasnova študije, ki vključuje poststratifikacijo in uteževanje, s katerim vzorce prilagodijo na podlagi podatkov iz Eurostata in na ta način omejijo pristranskost zaradi neodgovora (Mohorko in drugi 2011, 4).

Spremenljivke, za katere bomo analizirali pristranskost, so enake kot pri Mohorko in drugih (2011, 5): *spol*, *starost*, *starost ob koncu izobraževanja*, *politična usmeritev levo-desno* (merjena na lestvici od 1 do 10, kjer 1 pomeni najbolj levo, 10 pa najbolj desno) in *zadovoljstvo z življenjem* (merjeno na lestvici od 1 do 4, kjer 1 pomeni najmanjše

zadovoljstvo, 4 pa največje). Uporabnike interneta identificiramo na podlagi spremenljivke *anketiranec ima doma internetno povezavo*. Vrednosti spremenljivk in morebitno rekodiranje so predstavljeni v prilogi A.

Vzorec Evrobarometra zajema prebivalce članic EU in držav kandidatk, ki so starejši od 15 let. Seznam vključenih držav ni popolnoma enak za vsa leta zaradi dogodkov, kot so pridružitve Hrvaške Evropski uniji leta 2013 in sprememb pri državah kandidatkah. Naša analiza se bo omejila na iste države kot Mohorko in drugi (2011): 27 držav, ki so bile članice EU v celotnem pokritem obdobju, in tri dodatne države – Hrvaško, Turčijo in turško skupnost na Cipru (TCC). Na tem mestu je treba omeniti še, da se podatki za Evrobarometer zbirajo ločeno za Vzhodno in Zahodno Nemčijo ter Veliko Britanijo in Severno Irsko. Če bi želeli analizirati Nemčijo ali Združeno kraljestvo kot celoto, je to mogoče z uporabo ustrezne uteži, vendar bomo v tem delu izbrali isti pristop kot Mohorko in drugi (2011) in vzorce obravnavali ločeno.

Podatki iz Evrobarometra vključujejo vnaprej pripravljene populacijske in poststratifikacijske uteži. Uporaba populacijskih uteži je obvezna pri deskriptivni statistični analizi skupin držav, saj so Evrobarometrični vzorci zelo podobnih velikosti (pri večini držav zajemajo okrog 1000 enot), ne glede na dejansko število prebivalcev po državah. Populacijske uteži zagotavljajo, da so države v izbrani skupini zastopane proporcionalno, glede na to kolikšen delež celotne populacije predstavljajo njihovi prebivalci. Poststratifikacijske uteži na nivoju posameznih držav pa se uporablja s ciljem zmanjšati pristranskost zaradi neodgovora tako, da se struktura vzorca prilagodi sociodemografskim značilnostim populacije (*Weighting overview, Eurobarometer Data Service*).

Pri analizi posameznih držav bomo uporabili poststratifikacijsko utež W1, pri analizi celotne skupine držav, kjer je treba uporabiti populacijsko utež, pa naletimo na nekaj težav. Za naš vzorec bi bila najbolj primerna utež W25, ki vključuje 27 članic EU, Hrvaško, Turčijo in TCC, vendar je ta utež na voljo samo za leti 2011 in 2012. Utež W16, ki vključuje vse zgornje države razen TCC, je na voljo za vsa izbrana leta razen za 2016. Podatki vključujejo tudi utež W23, ki zajema trenutnih 28 članic EU, torej brez Turčije in TCC. Da bi ohranili primerljivost z raziskavo Mohorko in drugih (2011), smo uporabili naslednje populacijske uteži:

- W25 za leti 2011 in 2012;

- W16 za leta 2013, 2014 in 2015, pri čemer enote iz TCC pač niso utežene. Natančneje, v datoteki podatkov je bila pri uteži W16 pri teh enotah manjkajoča vrednost in le-to smo rekodirali v 1, kar pomeni, da teh enot ne utežimo;
- W23 za leto 2016, kjer enote iz Turčije in TCC niso utežene. Tudi v tem primeru smo pri uteži W23, ki ima za ti dve območji manjkajočo vrednost, to vrednost rekodirali v 1, kar pomeni, da enote iz Turčije in TCC niso utežene.

Ta kompromis pomeni, da so podatki za Turčijo in TCC uteženi le za leti 2011 in 2012, za ostala leta pa ne. Primerjava relativne pristranskosti (za postopek izračuna glej poglavje 6.3) za pet izbranih spremenljivk za leto 2011 v primeru, da so utežene enote iz vseh držav (uporabljena utež W25), da so utežene vse enote razen TCC (uporabljena utež W16 z vrednostjo 1 za TCC) ter da so utežene vse enote razen Turčije in TCC (uporabljena utež W23 z vrednostjo 1 za ti dve državi), pokaže, da je vpliv na izračune pristranskosti minimalen (gl. Tabelo 6.1). V vsakem primeru pa je tak kompromis edina možnost, ki jo imamo, ker populacijske uteži za vsa leta za vse države pač niso na razpolago.

Tabela 6.1 Primerjave relativne pristranskosti za leto 2011 glede na uporabljenno utež

RELATIVNA PRISTRANSKOST (v%)					
Utež	Spol	Starost	Izobrazba	Levo-desno	Zadovoljstvo
W23 + (TUR=1)+ (TCC=1)	-3,6 %	-10,3 %	6,1 %	-0,7 %	4,3 %
W16 + (TCC=1)	-3,5 %	-9,9 %	6,8 %	-1,8 %	4,5 %
W25	-3,6 %	-9,9 %	6,8 %	-1,7 %	4,5 %

V enem letu se izvaja več valov Evrobarometra, ki se osredotočajo na različne tematike in se razlikujejo po zastavljenih vprašanjih. Za analizo smo izbrali zadnji val raziskave v vsakem letu, ki je vključeval vse izbrane spremenljivke. Če takšnega vala ni bilo na voljo, smo podatke za spremenljivke pridobili iz različnih valov v istem letu. Kljub temu ostaja nekaj manjkajočih vrednosti. Za leto 2010 ni na voljo podatkov za TCC in podatka za zadovoljstvo z življenjem za Hrvaško in Turčijo. Za leto 2012 manjkajo podatki o politični usmeritvi za Hrvaško, Turčijo in TCC, medtem ko za leto 2013 manjkajo podatki o politični usmeritvi za Turčijo in TCC.

6.3 Metodologija

Za primerjavo uporabnikov in neuporabnikov interneta bomo uporabili bivariatne statistične analize. Pri spremenljivki spol, ki je nominalnega tipa, bomo uporabili test hi-kvadrat za

značilnost razlik v porazdelitvi med uporabniki in neuporabniki interneta. Pri ostalih demografskih in sociopolitičnih spremenljivkah, ki so intervalnega ali razmernostnega tipa, bomo uporabili t-test za neodvisna vzorca za značilnost razlik v povprečjih med uporabniki in neuporabniki interneta. Te analize bomo izvedli samo za Slovenijo in skupni vzorec držav EU, Turčije in TCC.

V drugem delu analize bomo izračunali pristranskost zaradi nepokritja po spremenljivkah ter naše ocene primerjali z ocenami Mohorko in drugih (2011). To analizo bomo izvedli tako za skupni vzorec držav EU, Turčije in TCC, kot tudi za posamezne države iz tega vzorca. Kot rečeno, je namen te analize ugotoviti, kakšno pristranskost zaradi nepokritja lahko pričakujemo, če v vzorec ne vključujemo neuporabnikov interneta.

Mohorko in drugi (2011, 7) uporabljajo dva indikatorja za merjenje pristranskosti zaradi nepokritja: relativno pristranskost in njeno absolutno vrednost. Relativno pristranskost uporabijo v deskriptivne namene, absolutno relativno pristranskost pa pri računanju razlik med državami in sprememb skozi leta, saj bi se drugače pozitivne in negativne vrednosti lahko med seboj izničile. Spodaj sta prikazani enačbi, ki ju uporabijo za izračun pristranskosti.⁶

$$1) \text{ relativna pristranskost zaradi nepokritja} = \frac{\bar{y}_{Int} - \bar{y}_{EB}}{\bar{y}_{EB}}$$

$$2) \text{ absolutna relativna pristranskost zaradi nepokritja} = \left| \frac{\bar{y}_{Int} - \bar{y}_{EB}}{\bar{y}_{EB}} \right|$$

EB predstavlja celotni vzorec iz Evrobarometra, ki ga obravnavamo kot ciljno populacijo, medtem ko *Int* predstavlja podvzorec uporabnikov interneta, ki ga obravnavamo kot pokrito populacijo pri potencialnih anketah iz (verjetnostnih) spletnih panelov, ki vključujejo le uporabnike interneta in zanemarijo neuporabnike. To pomeni, da $\bar{y}_{EB}$ in $\bar{y}_{Int}$ predstavljata povprečji določene spremenljivke *y* za celotni vzorec Evrobarometra in podvzorec uporabnikov interneta. Razlika med tema povprečjema je absolutna pristranskost, ki jo

⁶ Pri enačbi 1 gre za preoblikovano različico enačbe za relativno pristranskost, ki smo jo predstavili v teoretičnem delu. Končni rezultat za relativno pristranskost je isti, ne glede na to, katero obliko enačbe uporabimo.

merimo na enaki merski lestvici kot izbrano spremenljivko⁷. Če absolutno pristranskost delimo s povprečjem za celoten vzorec, dobimo relativno pristranskost. Pozitivne vrednosti relativne pristranskosti zaradi nepokritja nakazujejo, da bodo ocene, neupoštevajoč uporabnike, za povprečje izbrane spremenljivke previsoke, medtem ko negativne vrednosti nakazujejo, da bodo te ocene prenizke.

6.4 Analiza in rezultati

6.4.1 Penetracija interneta

Analizo podatkov iz Evrobarometra začnemo s pregledom frekvenc za spremenljivko *anketiranec ima doma internetno povezavo*, ki jo bomo uporabili za ločevanje med uporabniki in neuporabniki interneta. Ta spremenljivka služi kot indikator penetracije interneta, vendar moramo upoštevati, da nekateri člani gospodinjstev brez internetne povezave lahko internet uporabljajo kje druge (npr. v službi, šoli, knjižnici) ali pa, predvsem v zadnjih analiziranih letih, utegnejo uporabljati mobilni internet (gl. Fuchs in Busse 2009). Glede slednjega iz vprašanja, ki je uporabljeno v Evrobarometru, ni jasno, ali anketiranci z »internetno povezavo doma« razumejo samo fiksno internetno povezavo ali tudi mobilno internetno povezavo. Kljub tem nejasnostim in pomanjkljivostim je za našo analizo tako merjena spremenljivka glede penetracije primerna, saj se v primeru verjetnostnih spletnih panelov, ki se uporabljajo za anketiranje splošne populacije, kot neuporabnike interneta običajno obravnava prav tiste, ki internetne povezave nimajo doma (ne glede na to, ali internet uporabljajo kje drugje).

Tabela 6.2: Frekvence in odstotki za spremenljivko *anketiranec ima doma internetno povezavo*

	2010		2011		2012		2013		2014		2015		2016	
	Frek.	%	Frek.	%	Frek.	%	Frek.	%	Frek.	%	Frek.	%	Frek.	%
NE	11482	40	10629	36	9977	34	8501	29	8941	30	8729	30	8191	28
DA	17192	60	18584	64	19160	66	20829	71	20547	70	20484	70	21054	72
Skupaj	28674	100	29213	100	29137	100	29330	100	29488	100	29214	100	29245	100

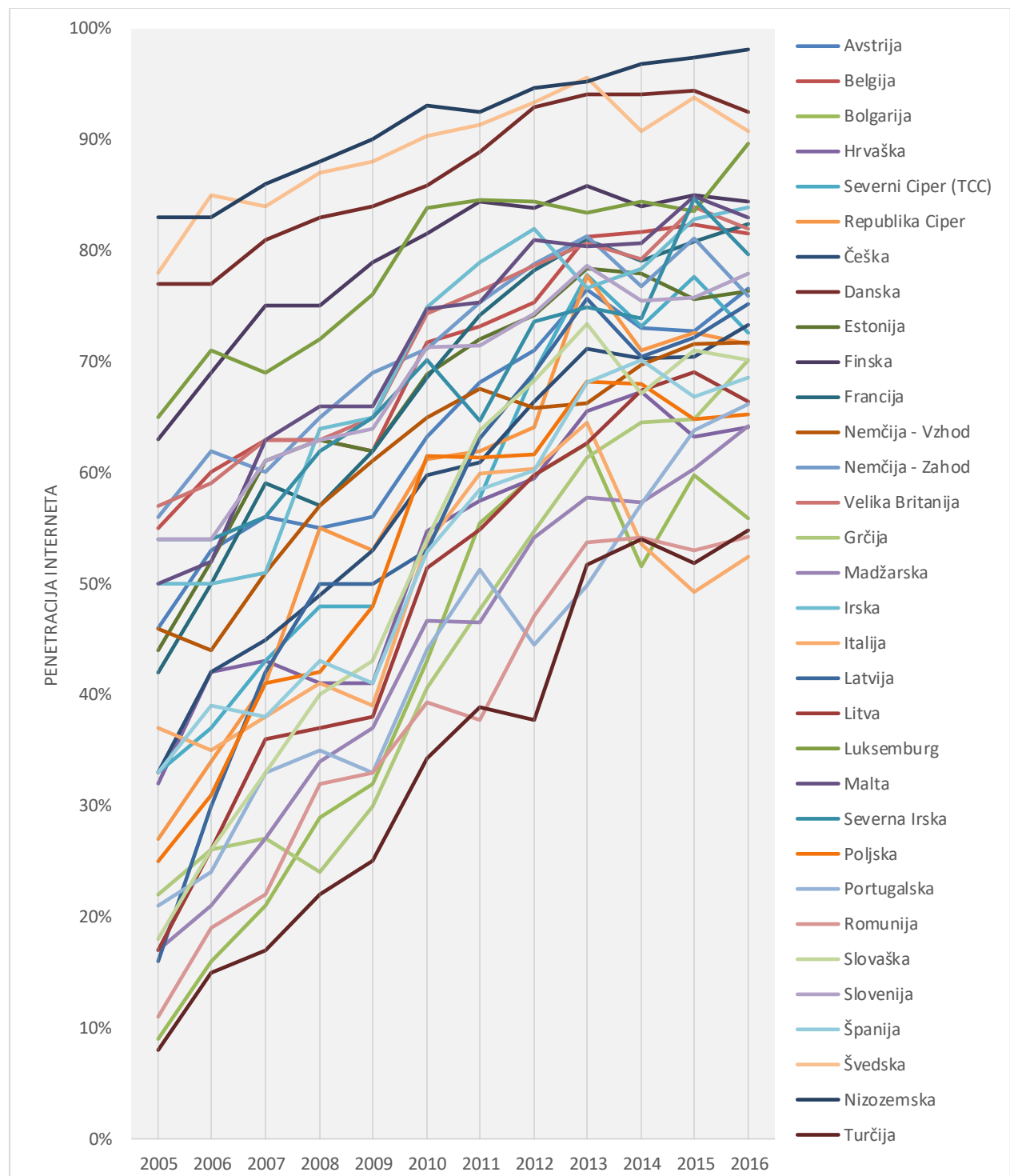
Tabela 6.2 prikazuje frekvence in deleže uporabnikov interneta za celoten vzorec, ki vključuje Evropsko unijo, Turčijo in TCC. Vidimo lahko, da je delež uporabnikov interneta od leta 2010 do leta 2016 narasel za 12 odstotnih točk. Večji del te rasti se je zgodil med leti

⁷ Absolutne pristranskosti ne smemo zamenjati z absolutno vrednostjo relativne pristranskosti. Relativno pristranskost definiramo v nadaljevanju, njena absolutna vrednost (označena z navpičnima oklepajema | |) pa govori o oddaljenosti te vrednosti od številskega izhodišča (točke 0), kar uporabimo, ko nas zanima le vrednost, ne pa tudi njen predznak.

2010 in 2013, medtem ko se za obdobje med leti 2013 in 2016 delež uporabnikov interneta giblje blizu vrednosti 70 %. Za primerjavo, delež uporabnikov od leta 2005 do leta 2009 je zrasel s 40 % na 55 % (Mohorko in drugi 2011, 23).

Slika 6.1 prikazuje rast dostopa do interneta po državah za obdobji 2005 do 2009 (gl. Mohorko in drugi 2011, 23) in 2010 do 2016 (gl. prilogo B). Vidimo lahko, da je v prvem obdobju precej bolj konsistenten trend rasti, medtem ko v drugem na splošno opazimo rast, na nivoju držav pa prihaja do bolj izrazitih nihanj. Zanimiv je primer Italije, kjer je delež dostopa do interneta od doma celo upadel s 53.6 % leta 2010 na 52.4 % leta 2016, medtem ko je bil na svoji najvišji točki leta 2016 64.5 %. To bi lahko deloma pojasnili s prehodom od stacionarnega k mobilnemu internetu.

Slika 6.1 Rast dostopa do interneta od doma med leti 2005 in 2016



6.4.2 Razlike med uporabniki in neuporabniki interneta

6.4.2.1 Celotni vzorec (EU, Turčija in TCC)

V Tabeli 6.3 so prikazani rezultati statističnih testov za razlike med uporabniki in neuporabniki interneta za celotni vzorec po posameznih letih.

Tabela 6.3: Bivariatne analize

	Uporabniki interneta	Neuporabniki interneta	Vrednost t-statistike ali χ^2 statistike	Natančna stopnja značilnosti t-testa ali χ^2 testa
2010				
SPOL			107,170	0,000***
Moški	51 %	45 %		
Ženski	49 %	55 %		
STAROST	41,2	53,2	55,9	0,000***
IZOBRAZBA	19,9	16,4	-61,8	0,000***
POLITIČNA USMERITEV	5,2	5,4	8,5	0,000***
ZADOVOLJSTVO	-	-	-	-
2011				
SPOL			68,172	0,000***
Moški	50 %	45 %		
Ženski	50 %	55 %		
STAROST	41,3	53,8	55,8	0,000***
IZOBRAZBA	19,3	16,0	-62,2	0,000***
POLITIČNA USMERITEV	5,3	5,6	8,4	0,000***
ZADOVOLJSTVO	3,1	2,7	-35,9	0,000***
2012				
SPOL			34,096	0,000***
moški	50 %	46 %		
ženski	50 %	54 %		
STAROST	41,5	54,1	54,7	0,000***
IZOBRAZBA	19,3	15,8	-66,7	0,000***
POLITIČNA USMERITEV	-	-	-	-
ZADOVOLJSTVO	3,0	2,7	-36,1	0,000***
2013				
SPOL			29,634	0,000***
moški	50 %	46 %		
ženski	50 %	54 %		
STAROST	41,4	56,4	63,3	0,000***
IZOBRAZBA	19,4	15,8	-66,7	0,000***
POLITIČNA USMERITEV	5,06	5,12	1,7	0,098*
ZADOVOLJSTVO	3,0	2,6	-32,9	0,000***

2014				
SPOL			152,283	0,000***
moški	51 %	43 %		
ženski	49 %	57 %		
STAROST	42,2	55,8	58,2	0,000***
IZOBRAZBA	19,6	16,3	-55,0	0,000***
POLITIČNA USMERITEV	5,2	5,3	4,1	0,000***
ZADOVOLJSTVO	3,0	2,7	-32,8	0,000***
2015				
SPOL			45,923	0,000***
moški	50 %	45 %		
ženski	50 %	55 %		
STAROST	42,6	57,6	63,6	0,000***
IZOBRAZBA	19,4	16,1	-57,4	0,000***
POLITIČNA USMERITEV	5,2	5,6	11,9	0,000***
ZADOVOLJSTVO	3,1	2,7	-31,8	0,000***
2016				
SPOL			39,827	0,000***
moški	50 %	46 %		
ženski	50 %	54 %		
STAROST	43,4	58,1	59,9	0,000***
IZOBRAZBA	19,7	16,6	-46,8	0,000***
POLITIČNA USMERITEV	5,1	5,3	4,4	0,000***
ZADOVOLJSTVO	3,1	2,8	-34,6	0,000***

Statistične značilnosti so prikazane tristopenjsko: *** $p < 0,01$; ** $0,01 \leq p < 0,05$; * $0,05 \leq p < 0,1$

Če povzamemo, v letih, ki jih pokriva naša analiza, ni prišlo do pomembnih sprememb v značilnostih uporabnikov in neuporabnikov interneta. Med uporabniki in neuporabniki ostajajo statistično značilne razlike, ki so posebej velike pri povprečni starosti, izstopajo pa tudi pri izobrazbi in sestavi glede na spol. Uporabniki interneta so v povprečju mlajši od neuporabnikov in so izobraževanje končali pri višji starosti. Leta 2016 je bila tako povprečna starost med uporabniki interneta 43 let, med neuporabniki pa 58 let, povprečna starost pri končani izobrazbi pa je bila med uporabniki skoraj 20 let, med neuporabniki pa le 16,6. Med uporabniki je vsa leta večji delež moških kot med neuporabniki. Leta 2016 je bil delež moških in žensk med uporabniki interneta enak, medtem ko so med neuporabniki prevladovalе ženske, ki jih je bilo 54 %. Razlike pri zadovoljstvu z življenjem in politični usmeritvi so manjše, vendar še vedno velja, da so uporabniki interneta v povprečju nekoliko bolj zadovoljni in se bolj nagibajo k politični levici kot neuporabniki. Leta 2016 je bilo povprečno zadovoljstvo (na lestvici od 1 do 4) uporabnikov interneta 3,1, neuporabnikov pa 2,8. Pri politični usmeritvi je bila v tem letu razlika med skupinama samo 0,2 točke na

lestvici od 1 do 10; povprečje uporabnikov interneta je bilo 5,1, neuporabnikov pa 5,3. Za vse spremenljivke je bila natančna stopnja značilnosti manjša od 0,01.

6.4.2.2 Slovenija

Oglejmo si še razlike med uporabniki in neuporabniki interneta samo na vzorcu prebivalcev Slovenije, ki so prikazane v Tabeli 6.4.

Tabela 6.4: Bivariatne analize za Slovenijo za leto 2010

	Uporabniki interneta	Neuporabniki interneta	Vrednost t-statistike ali χ^2 statistike	Natančna stopnja značilnosti t-testa ali χ^2 testa
2010				
SPOL			4,631	0,036**
Moški	51 %	44 %		
Ženski	49 %	56 %		
STAROST	39,80	61,37	20,04	0,000***
IZOBRAZBA	20,29	17,47	-11,58	0,000***
POLITIČNA USMERITEV	5,32	5,24	-0,36	0,718
ZADOVOLJSTVO	3,17	2,94	-5,26	0,000***
2011				
SPOL			4,855	0,031**
moški	51 %	44 %		
ženski	49 %	56 %		
STAROST	40,24	62,47	20,98	0,000***
IZOBRAZBA	19,99	17,28	-11,58	0,000***
POLITIČNA USMERITEV	5,33	5,27	-0,25	0,799
ZADOVOLJSTVO	3,14	2,78	-7,41	0,000***
2012				
SPOL			1,46	0,252
moški	50 %	46 %		
ženski	50 %	54 %		
STAROST	40,75	63,63	21,08	0,000***
IZOBRAZBA	20,10	17,51	-11,18	0,000***
POLITIČNA USMERITEV	5,06	5,30	0,76	0,447
ZADOVOLJSTVO	3,10	2,92	-4,00	0,000***
2013				
SPOL			6,557	0,012**
moški	51 %	42 %		
ženski	49 %	58 %		
STAROST	41,45	65,12	21,172	0,000***
IZOBRAZBA	20,04	17,45	-9,156	0,000***
POLITIČNA USMERITEV	5,18	5,36	0,824	0,410
ZADOVOLJSTVO	2,96	2,72	-4,264	0,000***
2014				
SPOL			5,164	0,026**
Moški	51 %	43 %		
Ženski	49 %	57 %		

STAROST	42,61	61,86	16,218	0,000***
IZOBRAZBA	20,17	18,09	-6,744	0,000***
POLITIČNA USMERITEV	4,63	4,93	1,293	0,196
ZADOVOLJSTVO	3,07	2,85	-4,354	0,000***
2015				
SPOL			7,990	0,005***
moški	51 %	41 %		
ženski	49 %	59 %		
STAROST	43,46	66,12	18,516	0,000***
IZOBRAZBA	20,33	18,14	-7,685	0,000***
POLITIČNA USMERITEV	4,67	4,92	1,038	0,300
ZADOVOLJSTVO	3,18	2,97	-4,485	0,000***
2016				
SPOL			7,990	0,005***
moški	51 %	41 %		
ženski	49 %	59 %		
STAROST	43,46	66,12	18,516	0,000***
IZOBRAZBA	20,33	18,14	-7,685	0,000***
POLITIČNA USMERITEV	4,67	4,92	1,038	0,300
ZADOVOLJSTVO	3,18	2,97	-4,485	0,000***

Razlike med uporabniki in neuporabniki interneta v Sloveniji so, tako kot v celotnem vzorcu, največje pri starosti. Te razlike so za Slovenijo celo večje kot za celotni vzorec, saj je povprečna starost neuporabnikov interneta kar precej višja. Leta 2016 je bila povprečna starost uporabnikov 45 let, neuporabnikov pa 63 let. Razlike v izobrazbi so podobne za vsa leta analize. Kot za celotni vzorec se je tudi za Slovenijo izkazalo, da so uporabniki interneta v povprečju starejši od neuporabnikov, ko končajo z rednim izobraževanjem. Povprečna starost pri končani izobrazbi leta 2016 je bila 20 let za uporabnike in 18 let za neuporabnike interneta. Porazdelitev po spolu je bila za vsa leta precej podobna pri uporabnikih interneta: moških je bilo med uporabniki bodisi 50 % bodisi 51 %. Pri neuporabnikih se je porazdelitev z leti nekoliko bolj spreminjala. V vseh letih so v tej skupini prevladovali ženske; največji delež žensk je bil 59 % leta 2015, najmanjši pa 54 % leta 2012, ko razlike po spolu niso bile statistično značilne. Po zadnjih podatkih iz leta 2016 je bilo med neuporabniki 42 % moških in 58 % žensk. Razlike v politični usmeritvi (na lestvici od 1 do 10) so se izkazale za značilne samo v letu 2016, ko se je pokazalo, da so uporabniki interneta bolj levo politično usmerjeni od neuporabnikov (povprečna vrednost za to spremenljivko je bila 4,7 za uporabnike in 4,9 za neuporabnike). Razlike v zadovoljstvu z življenjem (na lestvici od 1 do 4) so bile precej konsistentne iz leta v leto; uporabniki interneta so nekoliko bolj zadovoljni od neuporabnikov. Leta 2016 je bilo povprečno zadovoljstvo uporabnikov 3,18, neuporabnikov pa 2,97.

6.4.3 Absolutna pristranskost zaradi nepokritja

V zadnjem delu analize predstavljamo, kakšne bi bile vrednosti pristranskosti zaradi nepokritja za izbrane spremenljivke, če bi v vzorec (npr. vzorec v okviru spletnega panela) vključili samo uporabnike interneta. Negativne vrednosti pristranskosti pomenijo, da bi bile ocene, pridobljene z anketo na takem vzorcu, prenizke, pozitivne vrednosti pa kažejo, da bi bile ocene previsoke. Najprej si ogledamo vrednosti absolutnih pristranskosti po spremenljivkah, nato pa se osredotočimo na relativno pristranskost, saj na ta način lahko primerjamo vrednosti pristranskosti za spremenljivke z različnimi merskimi lestvicami. Celotne tabele relativne in absolutne pristranskosti po državah so predstavljene v prilogi C.

Tabela 6.5 prikazuje vrednosti absolutnih pristranskosti za celoten vzorec. Opazimo, da ima pristranskost negativne vrednosti pri spremenljivkah spol, starost in politična usmeritev, pozitivne pa pri spremenljivkah starost pri končani izobrazbi in zadovoljstvo z življenjem. To pomeni, da bi z anketo, ki bi vključevala le uporabnike interneta, podcenili delež žensk, povprečno starost in desno politično usmeritev ter precenili starost pri končani izobrazbi in zadovoljstvo z življenjem. Pri spremenljivki starost so vrednosti precej visoke in kažejo, da bi zaradi nepokritja za več kot 4 leta podcenili povprečno starost v splošni populaciji. Starost pri končani izobrazbi bi precenili v povprečju za nekoliko več kot 1 leto. Pri politični usmeritvi, ki je merjena na lestvici od 1 (levo) do 10 (desno), so vrednosti absolutne pristranskosti precej majhne, kar pomeni, da bi s spletno anketo le rahlo precenili nagnjenost k levici. Zadovoljstvo z življenjem je merjeno na lestvici od 1 (sploh ni zadovoljen/zadovoljna) do 4 (zelo zadovoljen/zadovoljna); večje vrednosti pristranskosti in manjši razpon merske lestvice kažejo, da bi se pri tej spremenljivki zaradi nepokritja zmotili bolj kot pri politični usmeritvi, in sicer bi precenili povprečno zadovoljstvo z življenjem.

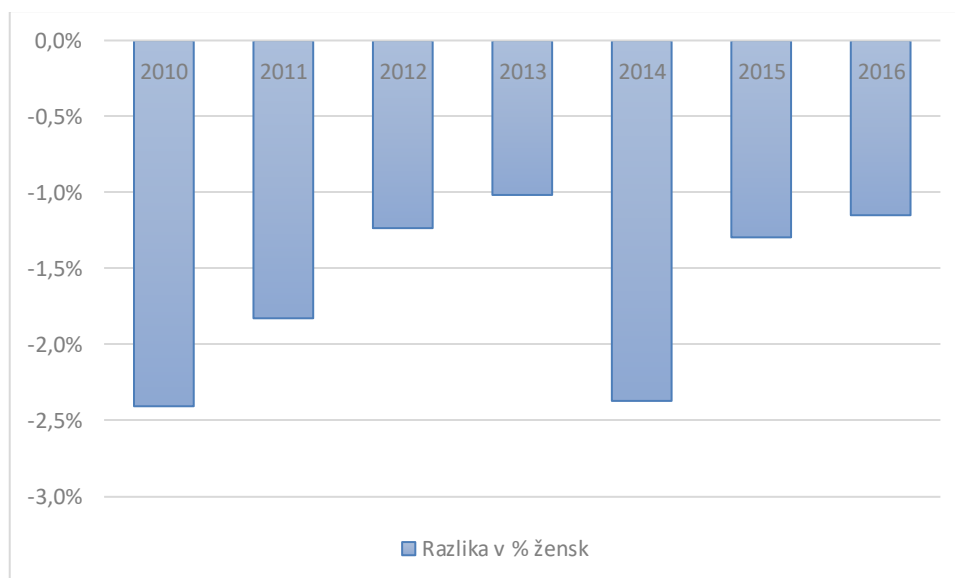
Tabela 6.5: Absolutna pristranskost po letih za vse spremenljivke

	Spol (% žensk)	Starost (leta)	Starost pri končani izobrazbi (leta)	Politična usmeritev (povprečne ocene na lestvici 1–10)	Zadovoljstvo z življenjem (povprečne ocene na lestvici 1–4)
2010*	-2,5 %	-4,938	1,313	-0,109	
2011	-1,8 %	-4,557	1,233	-0,090	0,131
2012	-1,2 %	-4,304	1,210		0,126
2013	-1,0 %	-4,344	1,101		0,103
2014	-2,4 %	-4,123	1,027	-0,038	0,103
2015	-1,3 %	-4,459	1,010	-0,109	0,097
2016	-1,1 %	-4,112	0,864	-0,037	0,095
Povprečje	-1.6 %	-4,405	1,108	-0,077	0,109

*vzorec za leto 2010 ne vključuje TCC

Oglejmo si še bolj podrobno, kako se absolutna pristranskost za spol, starost, starost pri končani izobrazbi, politično usmeritev in zadovoljstvo z življenjem spreminja med leti 2010 in 2016.

Slika 6.2 Absolutna pristranskost zaradi nepokritja pri spremenljivki spol

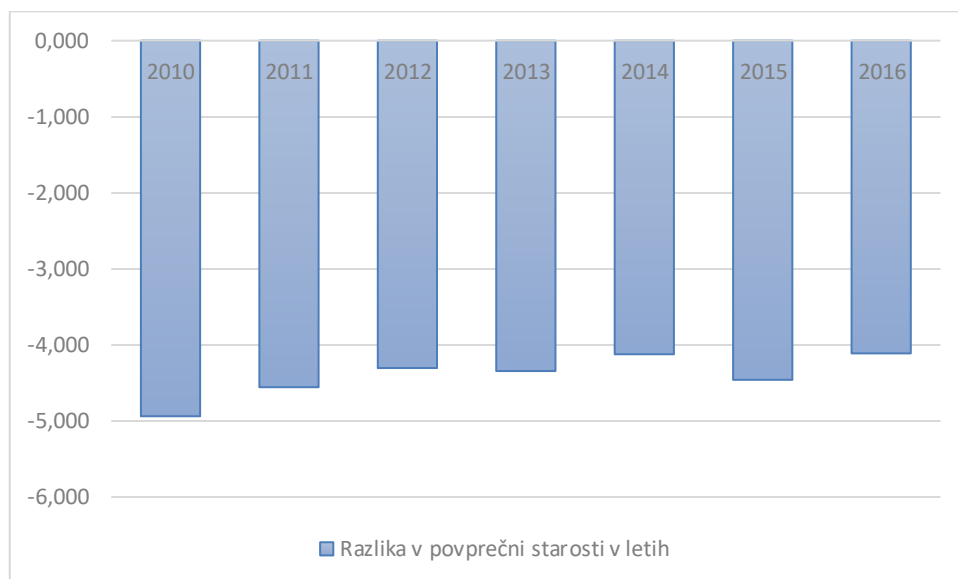


Slika 6.2 prikazuje vrednosti absolutne pristranskosti po letih za spol. Med leti 2010 in 2013 opazimo postopno upadanje pristranskosti, leta 2014 pa sledi nepričakovan dvig. Vrednosti za leti 2015 in 2016 sta ponovno nižji, vendar ne dosežeta najmanjše izmerjene pristranskosti iz leta 2013. Ker pa gre v vseh letih za precej majhne vrednosti pristranskosti (razlika v

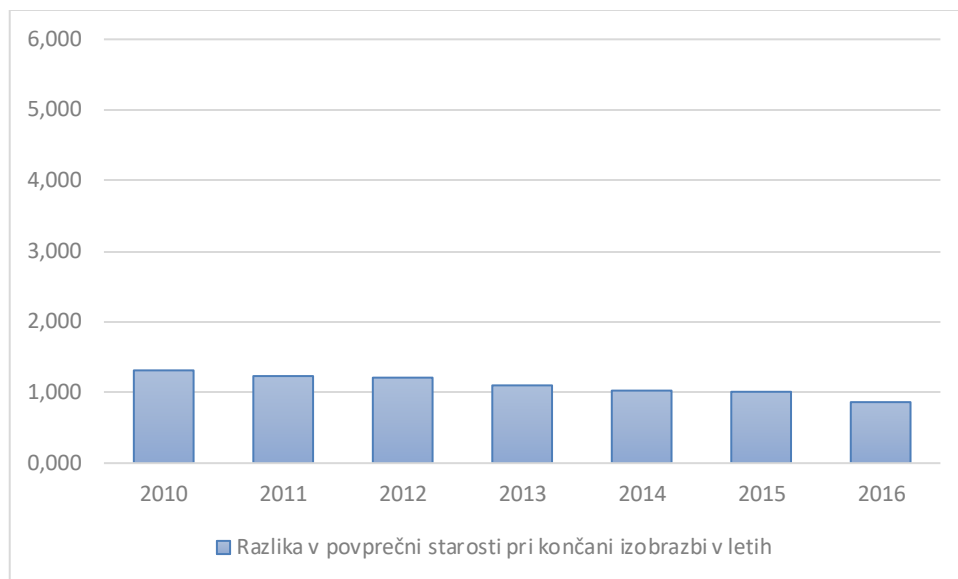
odstotku žensk na celotnem vzorcu in le na vzorcu uporabnikov sega od 1 do 2,5 odstotne točke), o nekih velikih nihanjih težko govorimo.

Slika 6.3 prikazuje vrednosti pristranskosti po letih za starost. Med leti 2010 in 2012 pristranskost upada, nato sledijo nihanja med leti 2013 in 2015, dokler ne doseže svoje najnižje vrednosti leta 2016. Tudi tu je razlika v povprečni starosti celotnega vzorca in vzorca uporabnikov interneta od 4,1 do 4,9 leta, kar pomeni, da gre za majhna nihanja med leti. Slika 6.4 prikazuje vrednosti pristranskosti za starost pri končani izobrazbi. Tudi tukaj lahko opazimo trend upadanja pristranskosti, vendar gre za majhne spremembe med leti, saj bi samo z vključitvijo uporabnikov interneta povprečno starost pri končani izobrazbi precenili le za od 0,9 do 1,3 leta.

Slika 6.3: Absolutna pristranskost zaradi nepokritja pri spremenljivki starost

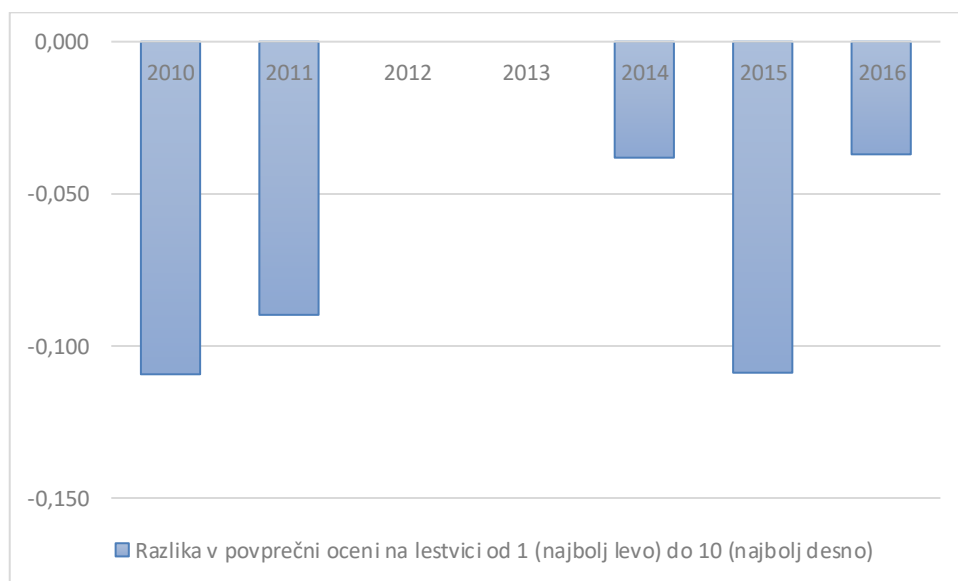


Slika 6.4 Absolutna pristranskost zaradi nepokritja pri spremenljivki starost pri končani izobrazbi

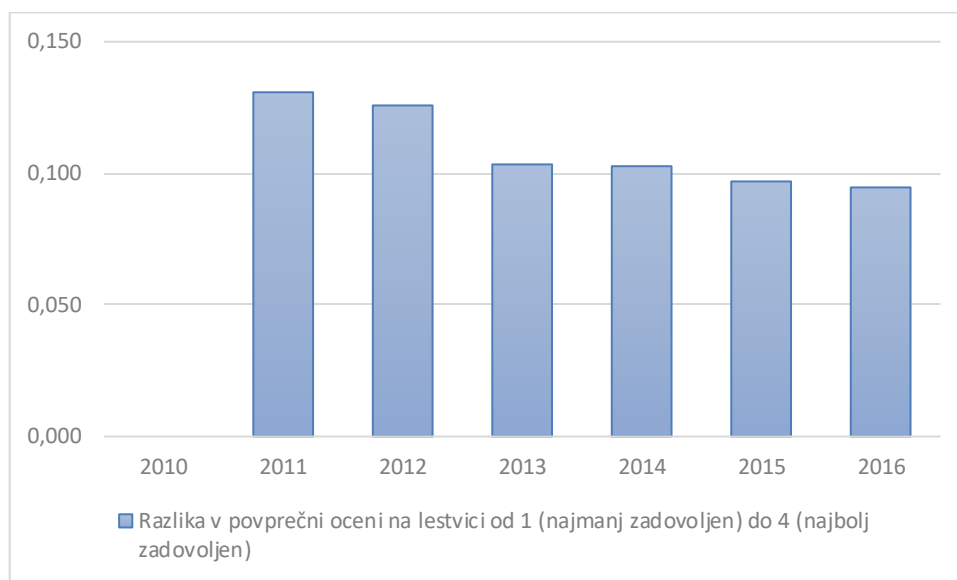


Slika 6.5 prikazuje vrednosti pristranskosti pri politični usmeritvi po letih. Pristranskost od leta 2010 do 2016 upade, opazimo pa lahko nihanja med posameznimi leti. Vendar so tudi tu vrednosti pristranskosti zelo nizke (le od 0,04 do 0,11 točke na lestvici od 1 do 10), kar dejansko ne pomeni bistvenih sprememb med leti. Pristranskost pri spremenljivki zadovoljstvo z življenjem od leta 2010 do leta 2016 konsistentno upada, največji padec pa lahko opazimo med letoma 2012 in 2013 (glej Sliko 6.6). Razlike zaradi pristranskosti so tudi tu precej majhne, med 0,10 in 0,13 točke na lestvici od 1 do 4, kar dejansko pomeni, da ne gre za bistvene spremembe med leti.

Slika 6.5 Absolutna pristranskost zaradi nepokritja pri spremenljivki politična usmeritev



Slika 6.6 Absolutna pristranskost zaradi nepokritja pri spremenljivki zadovoljstvo z življenjem



Če povzamemo, za večino spremenljivk med leti sicer prihaja do zmanjšanja absolutne pristranskosti zaradi nepokritja, vendar ne vedno linearno v času, hkrati pa so razlike v absolutni pristranskosti med leti minimalne.

6.4.4 Relativna pristranskost in primerjava z leti 2005 do 2009

Če želimo primerjati pristranskost pri spremenljivkah z različnimi merskimi lestvicami, je bolj primerno uporabiti relativno pristranskost, ki jo v Tabeli 6.6 izrazimo v odstotkih. Poleg let, vključenih v našo analizo, so tu prikazani še izračuni pristranskosti za leta 2005 do 2009 iz analize Mohorko in drugih (2011, 24–28). Primerjava podatkov za leto 2010 s podatki za leto 2016 pokaže padec relativne pristranskosti zaradi nepokritja pri vseh spremenljivkah. Po zadnjih podatkih iz leta 2016 je pristranskost zaradi nepokritja v spletnih anketah največja pri spremenljivki starost (za 8,65 % podcenimo povprečno starost), ki ji sledita spremenljivki starost pri končani izobrazbi (za 4,59 % precenimo povprečno starost pri končani izobrazbi) in zadovoljstvo z življenjem (za 3,15 % precenimo povprečno zadovoljstvo z življenjem), najmanjša pa pri spremenljivkah spol (za 2,23 % podcenimo delež žensk) in politična usmeritev (za 0,72 % podcenimo desno politično usmeritev). Manjše vrednosti pristranskosti pri demografskih spremenljivkah kažejo na krčenje digitalnega razkoraka, ki ga lahko povežemo s čedalje višjo penetracijo interneta.

Tabela 6.6: Relativna pristranskost po letih

	Spol	Starost	Starost pri končani izobrazbi	Politična usmeritev	Zadovoljstvo z življenjem
2005**	-2,70 %	-15,90 %	10,40 %	1,20 %	5,70 %
2006**	-2,90 %	-15,90 %	9,40 %	0,60 %	5,00 %
2007**	-2,00 %	-14,70 %	8,10 %	0,50 %	4,70 %
2008**	-1,80 %	-14,30 %	7,50 %	-0,50 %	4,60 %
2009**	-1,60 %	-13,60 %	7,40 %	-0,80 %	4,20 %
2010*	-4,67 %	-10,87 %	7,28 %	-2,08 %	-
2011	-3,55 %	-9,94 %	6,84 %	-1,67 %	4,47 %
2012	-2,40 %	-9,40 %	6,71 %	-	4,34 %
2013	-1,98 %	-9,50 %	6,02 %	-	3,61 %
2014	-4,61 %	-8,91 %	5,53 %	-0,73 %	3,50 %
2015	-2,51 %	-9,47 %	5,48 %	-2,07 %	3,26 %
2016	-2,23 %	-8,65 %	4,59 %	-0,72 %	3,15 %
Povprečje	-2,75 %	-11,76 %	7,10 %	-0,63 %	4,23 %

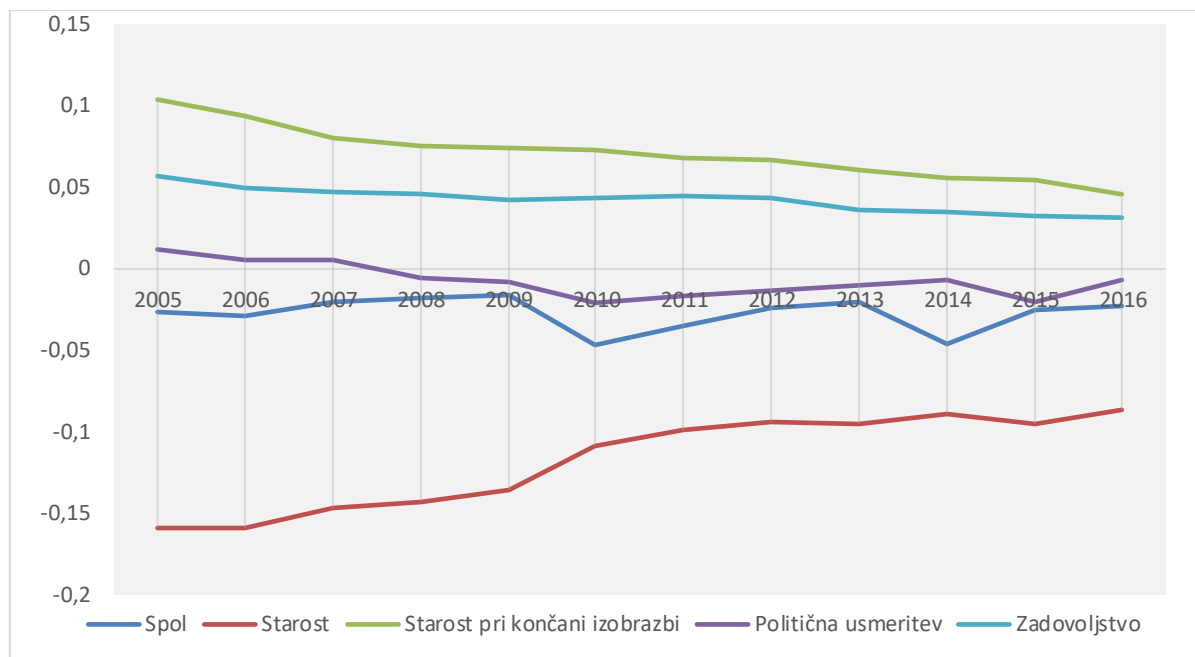
* podatki ne vključujejo TCC

**podatki so bili pridobljeni iz analize Mohorko in drugih (2011)

Slika 6.7 prikazuje spremembe pristranskosti po letih. Opazimo lahko upadanje pristranskosti (bližanje vrednosti 0 %) pri spremenljivkah starost, starost pri končani izobrazbi in zadovoljstvo z življenjem, ter nihanja pri spremenljivkah spol in politična usmeritev, kjer pa je pristranskost za vsa leta najnižja. Takšen rezultat v kontekstu večanja internetne penetracije kaže, da pristranskost konsistentno upada, dokler ne doseže dovolj nizke vrednosti, nato pa ostaja na podobnem nivoju in med leti celo niha. Pri zelo nizki pristranskosti, kot jo lahko opazimo pri spremenljivki politična usmeritev, se vrednosti z leti celo spreminjajo iz pozitivnih v negativne ali obratno. Glede na majhne razlike med uporabniki in neuporabniki interneta pri tej spremenljivki, ki smo jih prikazali v poglavju 6.4.2, lahko rečemo, da uporaba vzorca, ki vključuje le uporabnike interneta, ne bi pomembno vplivala na pristranskost pri oceni politične usmeritve za celotno Evropo. Ob tem velja opozoriti, da je pristranskost ocen politične usmeritve še vedno lahko večja na nivoju posameznih držav, o čemer bo govora v naslednjem podpoglavju.

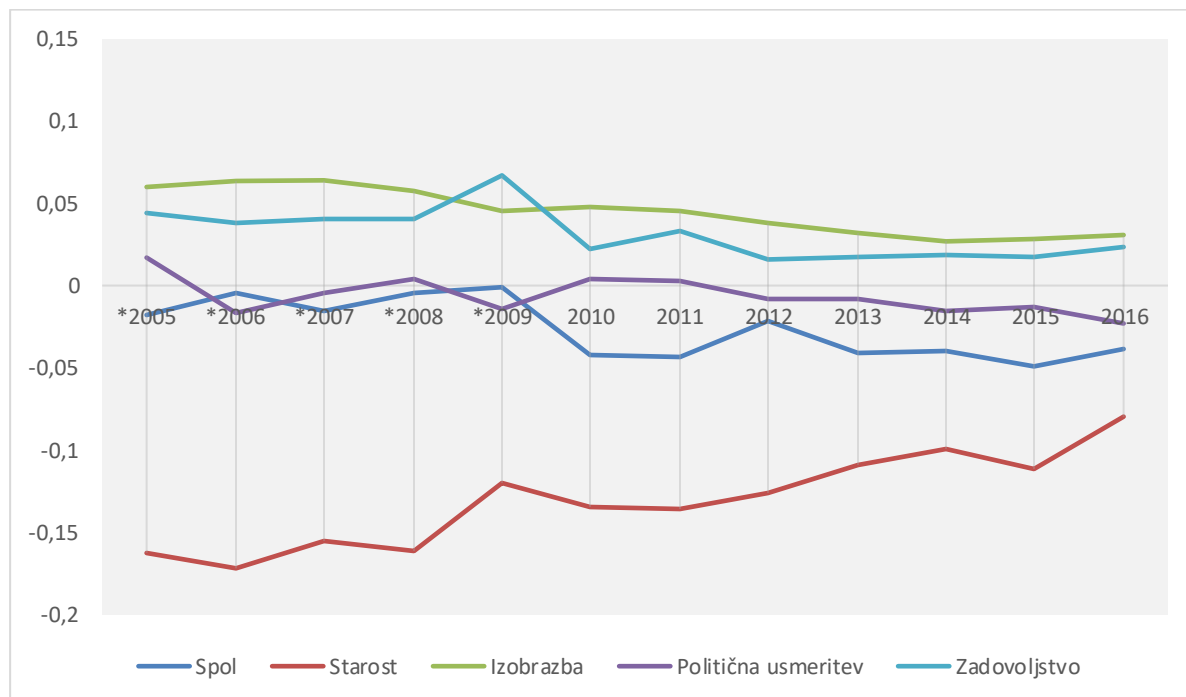
Rezultati analize za leta 2010 do 2016 se dokaj dobro ujemajo z rezultati Mohorko in drugih, saj se je pristranskost tudi v tem obdobju zmanjšala za vse spremenljivke. Kljub temu lahko v drugem obdobju opazimo nekaj posebnosti. Pristranskost za spremenljivko spol tako ne sledi pričakovanemu trendu manjšanja pristranskosti, ampak naraste na višje vrednosti, kot jih je imela v prvem obdobju, in začne med leti bolj nihati.

Slika 6.7: Relativna pristranskost za vse spremenljivke po letih



Slika 6.8 prikazuje spremembe v relativni pristranskosti po letih za Slovenijo, kjer opazimo podobne trende za večino spremenljivk. Upad pristranskosti pri izobrazbeni spremenljivki je manj strm, saj je bila že od začetka nižja od evropskega povprečja, vidimo pa tudi več nihanj pri spremenljivki zadovoljstvo z življenjem. O relativni pristranskosti za Slovenijo po posameznih spremenljivkah bomo govorili v naslednjem podpoglavju.

Slika 6.8: Vrednosti relativne pristranskosti za Slovenijo



6.4.5 Analiza relativne pristranskosti po državah

V tem delu predstavljamo podrobnejšo analizo relativne pristranskosti pri državah, pri čemer za posamezno državo najprej predstavimo povprečno relativno pristranskost preko vseh analiziranih let in vseh spremenljivk ter to povežemo tudi s stopnjo penetracije interneta. V drugem delu pa za izbrane države predstavimo še spremembe v relativni pristranskosti skozi čas.

V Tabeli 6.7 so države, ki so bile vključene v analizo, razvrščene glede na povprečne vrednosti relativne pristranskosti, neupoštevajoč predznak (torej vrednosti absolutne relativne pristranskosti), za vsa leta med 2010 in 2016 in spremenljivke. Za primerjavo je dodana še povprečna relativna pristranskost za leto 2016, torej stanje po zadnjih podatkih. Absolutne vrednosti relativne pristranskosti smo uporabili, da se pozitivne in negativne vrednosti ne bi izničile med seboj, tako da lahko v tem primeru govorimo le o povprečni višini pristranskosti ocen, ne pa tudi o smeri pristranskosti. To je vseeno relevantno, saj je v večini primerov smer pristranskosti nespremenjena, z izjemo tistih primerov, ko je relativna pristranskost zadosti blizu 0.

Vidimo, da je povprečna pristranskost ocen v analiziranem obdobju na Nizozemskem le 0,8 %, medtem ko v Bolgariji doseže 10 %. V Bolgariji in Estoniji in na Švedskem je bila pristranskost za leto 2016 celo večja od povprečne pristranskosti, vendar samo v primeru Bolgarije lahko opazimo trend naraščanja pristranskosti skozi celotno obdobje (vzrok za to je predvsem rast pristranskosti pri spremenljivki politična usmeritev, kar je razvidno iz izpisa v Prilogi C.4). Vrednosti povprečne pristranskosti za Slovenijo so le nekoliko nižje kot pri vzorcu EU 27.

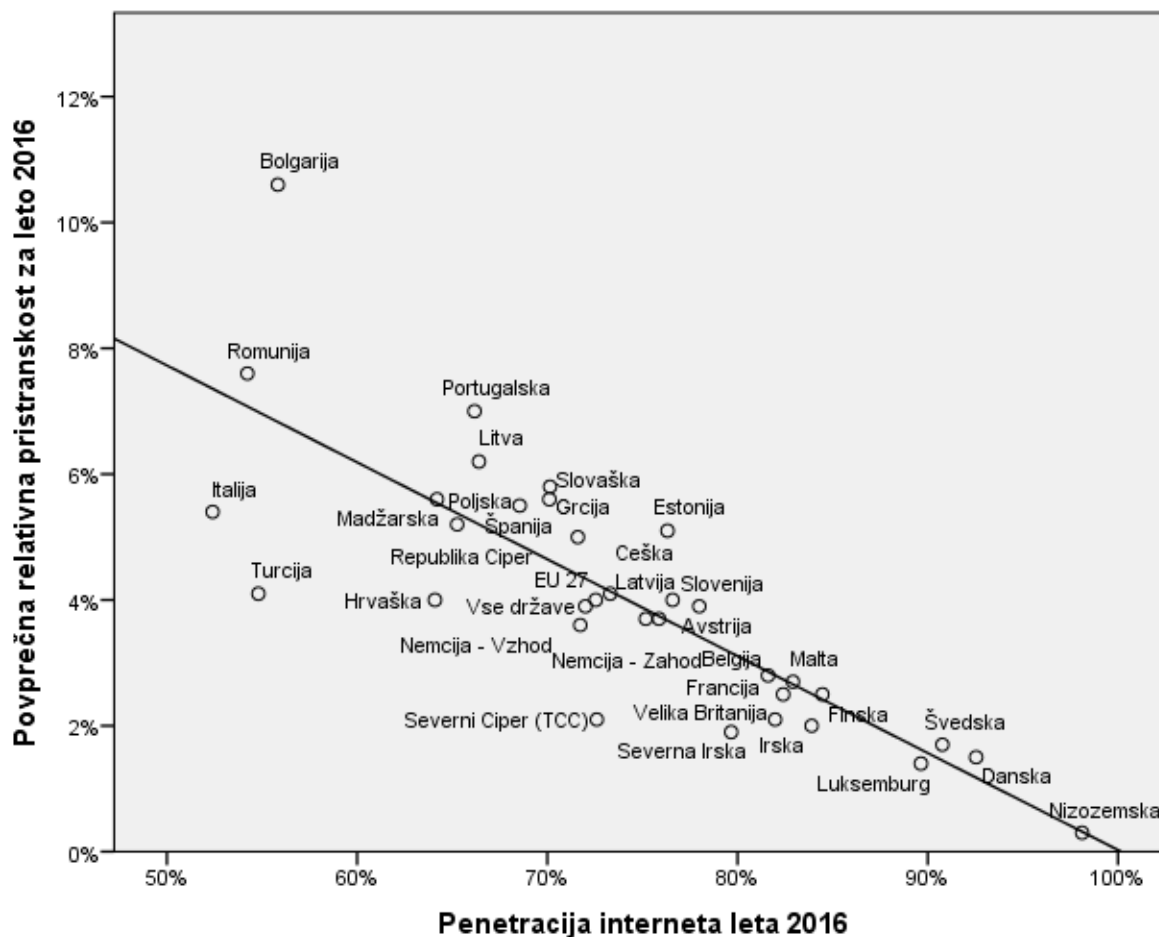
Tabela 6.7: Države glede na povprečno relativno pristranskost*

Država	Povprečna pristranskost za leta 2010 do 2016	Povprečna pristranskost za leto 2016	Penetracija interneta leta 2016
Nizozemska	0,8 %	0,3 %	98,1 %
Švedska	1,5 %	1,7 %	90,8 %
Danska	1,5 %	1,5 %	92,5 %
Luksemburg	2,6 %	1,4 %	89,6 %
Irska	2,6 %	2,0 %	83,9 %
Velika Britanija	2,8 %	2,1 %	82,0 %
Finska	3,0 %	2,5 %	84,5 %
Belgija	3,3 %	2,8 %	81,6 %
Severna Irska	3,5 %	1,9 %	79,7 %
Francija	3,6 %	2,5 %	82,4 %
Latvija	3,9 %	3,7 %	75,2 %
Malta	3,9 %	2,7 %	82,9 %
Nemčija - Zahod	4,0 %	3,7 %	75,9 %
Ciper – TCC	4,3 %	2,1 %	72,6 %
Slovenija	4,4 %	3,9 %	78,0 %
<i>EU 27</i>	4,6 %	4,0 %	72,5 %
Avstrija	4,6 %	4,0 %	76,6 %
Estonija	4,7 %	5,1 %	76,3 %
Nemčija - Vzhod	4,8 %	3,6 %	71,7 %
<i>Vse države</i>	4,9 %	3,9 %	72,0 %
Hrvaška	5,2 %	4,0 %	64,1 %
Poljska	5,4 %	5,2 %	65,3 %
Slovaška	5,7 %	5,6 %	70,1 %
Italija	5,8 %	5,4 %	52,4 %
Republika Ciper	6,0 %	5,0 %	71,6 %
Češka	6,0 %	4,1 %	73,3 %
Madžarska	6,3 %	5,6 %	64,2 %
Španija	6,4 %	5,5 %	68,5 %
Litva	6,4 %	6,2 %	66,4 %
Turčija	6,5 %	4,1 %	54,8 %
Romunija	8,3 %	7,6 %	54,2 %
Portugalska	8,3 %	7,0 %	66,2 %
Grčija	8,8 %	5,8 %	70,1 %
Bolgarija	10,0 %	10,6 %	55,8 %

*upoštevane so vrednosti relativne pristranskosti brez predznaka

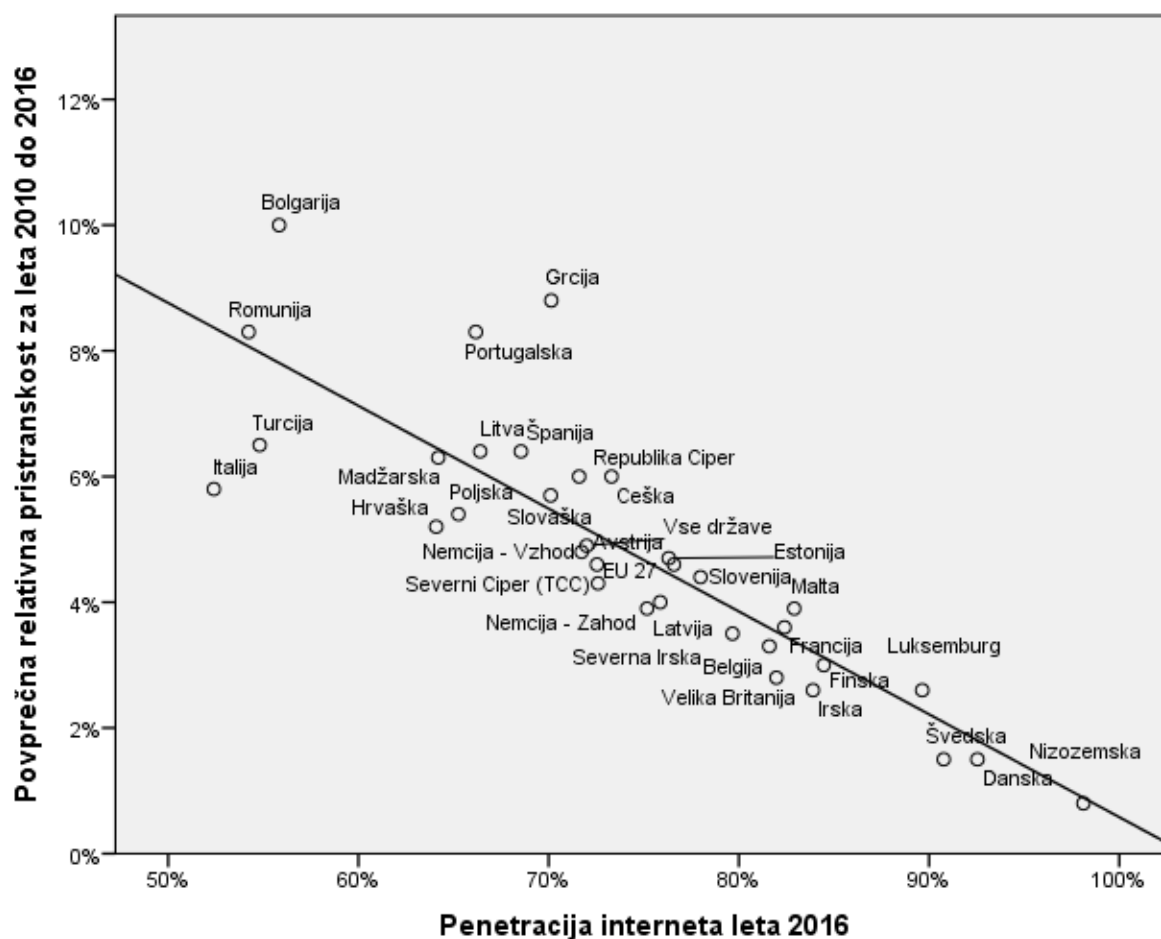
V Tabeli 6.7 lahko opazimo tudi, da imajo države z najnižjo povprečno pristranskostjo praviloma zelo visoko stopnjo penetracije interneta, zato si poglejmo korelacijo med stopnjo pristranskosti in odstotkom uporabnikov interneta v posamezni državi.

Slika 6.9: Povprečna (absolutna) relativna pristranskost za leto 2016 za analizirane države



Razsevni grafikon na Sliki 6.9 prikazuje odnos med stopnjo penetracije interneta leta 2016 in povprečno pristranskostjo za to leto. Razberemo lahko, da med tema spremenljivkama obstaja jasna negativna linearna povezanost; večja je penetracija interneta, manjša je povprečna pristranskost. Vrednost Pearsonovega koeficienta korelacije je -0.819, torej gre za močno povezanost. Močno negativno korelacijo opazimo tudi med penetracijo interneta leta 2016 in povprečno pristranskostjo za vsa leta (gl. Sliko 6.10). Vrednost Pearsonovega koeficienta korelacije je v tem primeru še višja (-0.859).

Slika 6.10: Povprečna (absolutna) relativna pristranskost za leta 2010 do 2016 za analizirane države



Ker je pristranskost povezana s stopnjo penetracije interneta, si oglejmo še razlike v relativni pristranskosti med državami glede na penetracijo interneta. V ta namen države razdelimo v pet skupin glede na stopnjo penetracije interneta za leto 2016, kot je prikazano v Tabeli 6.8. Za vsako skupino držav smo izračunali povprečno relativno pristranskost za posamezne spremenljivke za vsa leta.

Tabela 6.8: Države po skupinah glede na stopnjo penetracije interneta leta 2016

Stopnja penetracije interneta	Države
50 %-60 %	Italija, Romunija, Turčija, Bolgarija
60 %-70 %	Hrvaška, Madžarska, Poljska, Portugalska, Litva, Španija
70 %-80 %	Slovaška, Grčija, Republika Ciper, Nemčija - Vzhod, TCC, Češka, Latvija, Nemčija - Zahod, Estonija, Avstrija, Slovenija, Severna Irska
80 %-90 %	Belgija, Velika Britanija, Francija, Malta, Irska, Finska, Luksemburg
90 %-100 %	Švedska, Danska, Nizozemska

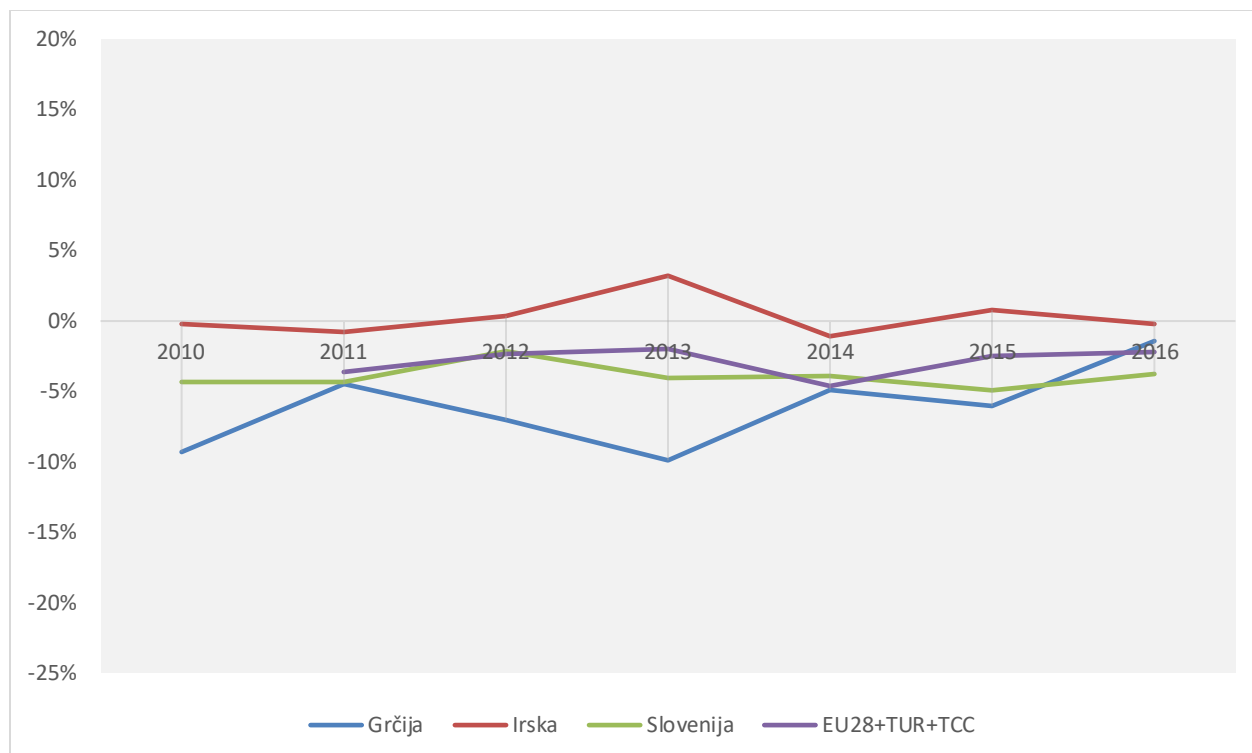
Tabela 6.9 kaže, da je povprečna relativna pristranskost pri vseh spremenljivkah najnižja v državah z visoko stopnjo penetracije. Morda je nekoliko nepričakovano, da je najvišja pristranskost pri spolu v skupini s penetracijo interneta med 70 % in 80 %. Države z najvišjo povprečno pristranskostjo pri spolu so Grčija (-6,1 %), Češka (-5,2 %), Vzhodna Nemčija (-5,1 %) in Slovaška (-6,1 %), ki se vse nahajajo ravno v tej srednji skupini po internetni penetraciji. Tudi pri starosti je povprečna relativna pristranskost najvišja v skupini s penetracijo 60 %-70 %, v skupini z najnižjo penetracijo pa je za 0,6 odstotne točke nižja. Razlog za tak rezultat je, da imata dve državi v skupini z najnižjo penetracijo (Turčija in Italija) stopnjo pristranskosti okrog 11 %, ki ni med najvišjimi. Razlike v pristranskosti zaradi nepokritja med skupinami držav glede na penetracijo interneta so nezanemarljive, vendar lahko vidimo, da stopnja penetracije interneta ni edini faktor, ki vpliva na pristranskost.

Tabela 6.9: Povprečna relativna pristranskost za leta 2010 do 2016 po skupinah držav glede na stopnjo penetracije interneta leta 2016

	Skupine držav glede na stopnjo penetracije interneta leta 2016				
	50 % - 60 %	60 % - 70 %	70 % - 80 %	80 % - 90 %	90 % - 100 %
SPOL	-3,0 %	-3,1 %	-3,9 %	-2,6 %	-1,4 %
STAROST	-14,3 %	-14,9 %	-11,7 %	-7,6 %	-3,2 %
IZOBRAZBA	8,1 %	6,8 %	4,3 %	3,1 %	1,0 %
POLITIČNA USMERITEV	3,5 %	-0,1 %	0,2 %	-0,3 %	-0,2 %
ZADOVOLJSTVO	5,3 %	4,6 %	2,6 %	1,0 %	0,3 %

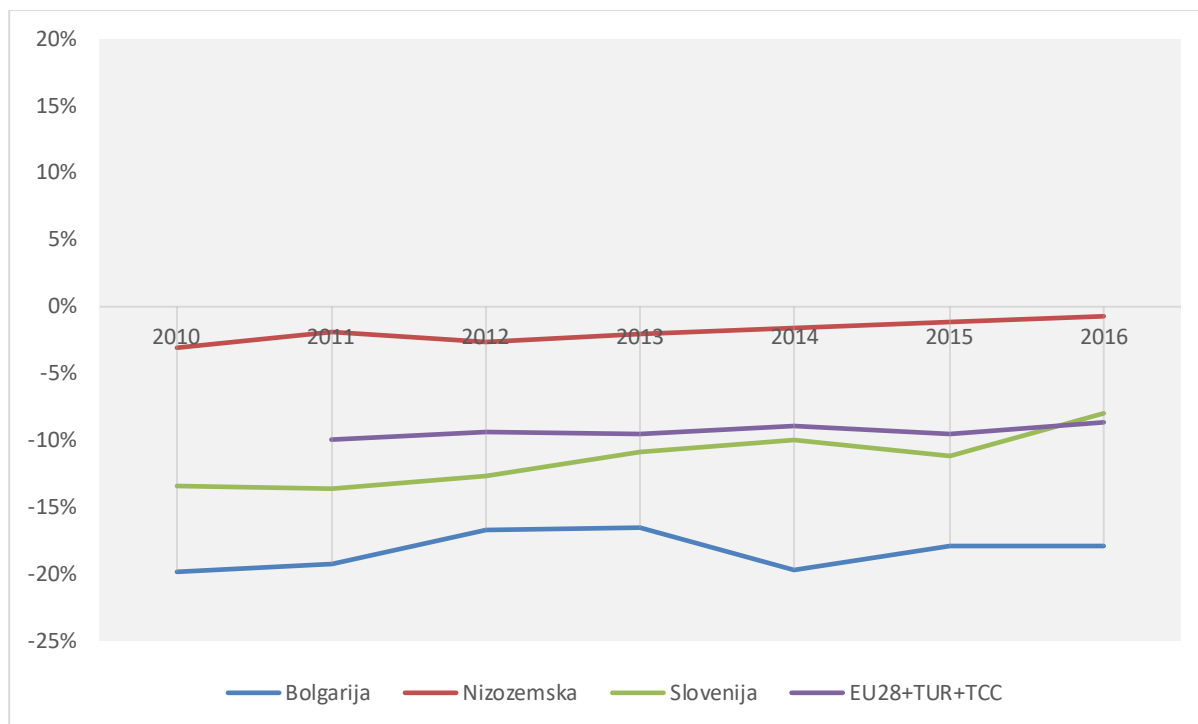
Oglejmo si še, do kakšnih razlik v relativni pristranskosti med leti lahko pride pri posameznih državah za analizirane spremenljivke. Zaradi boljše preglednosti v grafe ne bomo vključili vseh držav v vzorcu, temveč se bomo omejili na Slovenijo ter državi z najnižjo in najvišjo povprečno relativno pristranskostjo za vsako spremenljivko. Za primerjavo pa dodajamo tudi vrednosti za vse države.

Slika 6.11: Relativna pristranskost za spol skozi čas



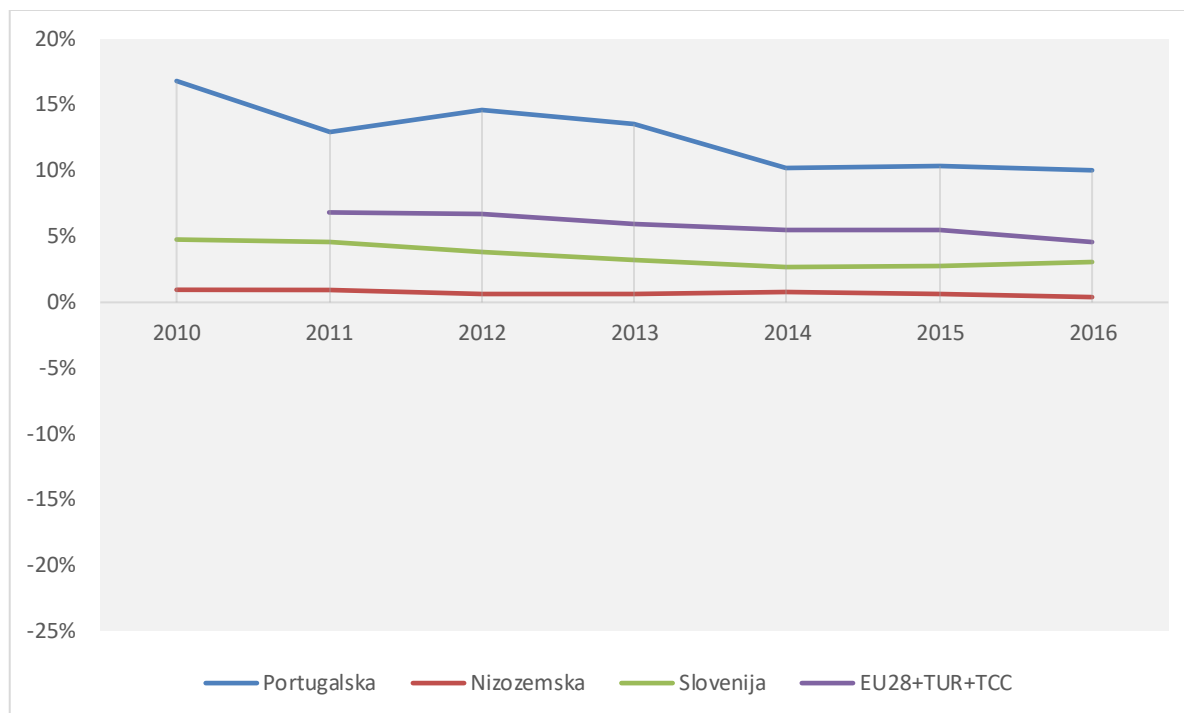
Pri spremenljivki spol (gl. Sliko 6.11) se pristranskost za Slovenijo ne razlikuje preveč od pristranskosti za celotni vzorec in je v obeh primerih negativna in precej majhna. Torej bi pri uporabi vzorca le uporabnikov interneta lahko pričakovali rahlo večji delež moških, ki pa v večini primerov verjetno ne bi pomembno vplival na zaključke raziskav, ki bi bile izvedene s pomočjo spletnih anket. Relativna pristranskost zaradi nepokritja pri spolu je v povprečju (preko vseh analiziranih let) najvišja v Grčiji, kjer leta 2013 doseže celo -10 % (absolutna pristranskost je v tem primeru -5 %), kar bi bilo vsekakor pomembno upoštevati pri izvajanju spletnih anket v tej državi. Kljub temu rezultati iz Grčije kažejo, da pristranskost pri spolu vztrajno upada, in za leta 2014 do 2016 je videti, kot da niha v podobnem razponu kot pri večini ostalih držav. Najnižja povprečna pristranskost zaradi spola je na Irskem, kjer za večino let ne doseže niti enega odstotka. Leto 2012, ko bi bile v hipotetični spletni anketi na Irskem preveč zastopane ženske, je izjema, ki kaže, da pri zelo nizkih vrednostih pristranskosti izstopajo tudi vrednosti, ki so same po sebi še vedno nizke.

Slika 6.12: Relativna pristranskost za starost skozi čas



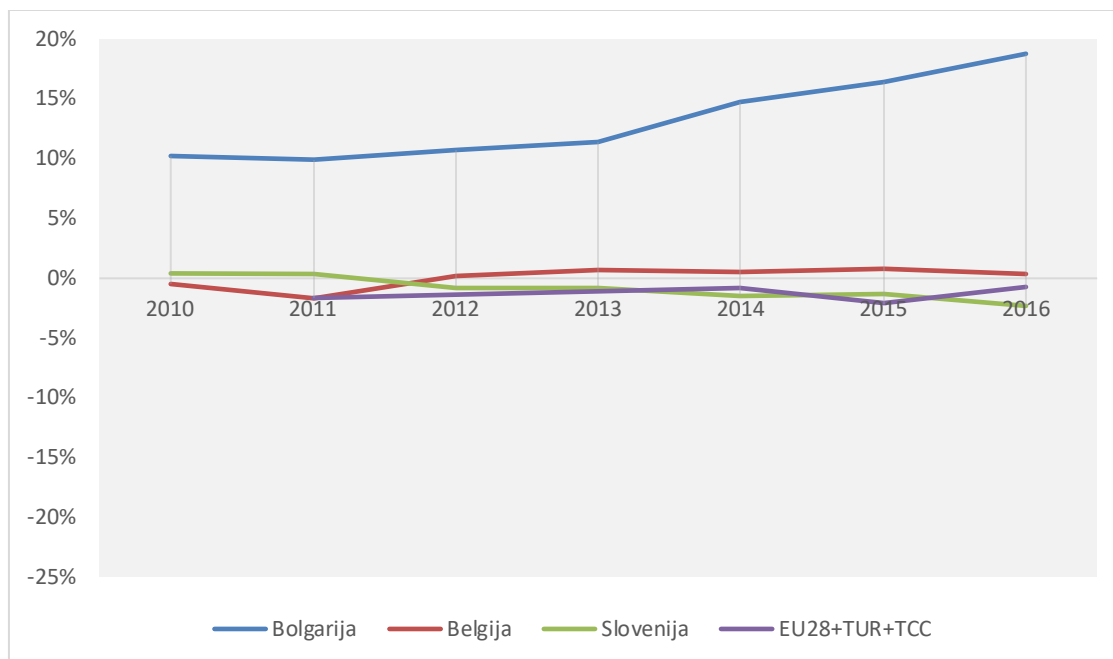
Na Sliki 6.12 lahko opazimo, da so razlike med državami za relativno pristranskost pri starosti precej večje kot pri spolu in da so vrednosti vedno negativne, kar pomeni, da bi spletne ankete v našem primeru vedno podcenile povprečno starost. Vrednosti pristranskosti za vse države upadajo zelo počasi, medtem ko se v Sloveniji to dogaja hitreje. Vrednosti relativne pristranskosti za Nizozemsko, kjer je povprečna relativna pristranskost pri starosti najnižja, so dosegle -1 %, kar pomeni, da v tej državi lahko pričakujemo, da uporaba spletnih anket ne bo vplivala na rezultate, kar se tiče starosti. V Sloveniji je vrednost relativne pristranskosti za leto 2016 okrog -8 % (v absolutnem smislu to pomeni -3,9 leta), kar kaže, da podatkov za raziskave, pri katerih je starost pomemben faktor, ni pametno zbirati samo prek spleta, čeprav lahko pričakujemo, da bo v naslednjih letih pristranskost še naprej upadala. Ta upad relativne pristranskosti je zelo počasen na nivoju celotnega vzorca, k čemur pripomorejo države, kot je Bolgarija, kjer je razlika v relativni pristranskosti med letoma 2010 in 2016 le 2 odstotni točki, v vmesnem obdobju pa vrednosti precej nihajo.

Slika 6.13: Relativna pristranskost za starost pri dokončani izobrazbi skozi čas



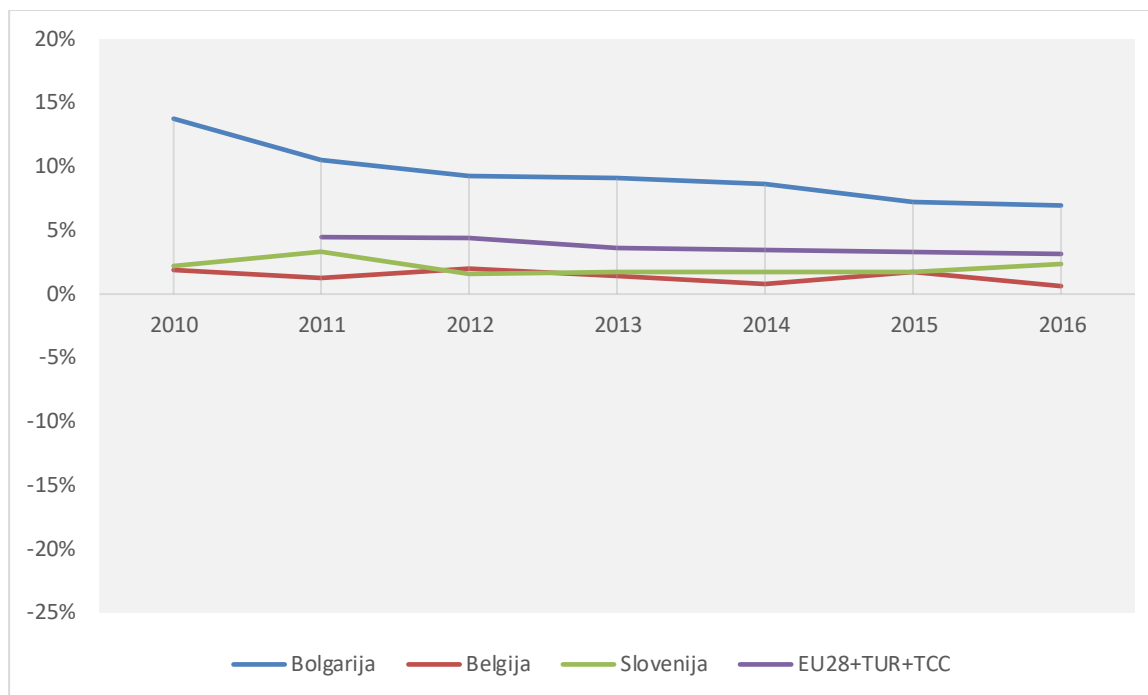
Relativna pristranskost za starost pri končani izobrazbi (gl. Sliko 6.13) ima v našem vzorcu vedno pozitivne vrednosti, kar pomeni, da bi jo s spletnimi anketami precenili. Razlike med državami so nekoliko manjše kot pri starosti, ampak še vedno dokaj velike. Vrednosti relativne pristranskosti za vse države postopoma upadajo. Leta 2016 dosežejo okoli 5 % za vse države, 3 % za Slovenijo in manj kot 1 % za Nizozemsko, kjer je ta pristranskost za to spremenljivko najnižja (vrednosti absolutne pristranskosti so 0,86 leta za vse države, 0,61 leta za Slovenijo in 0,08 za Nizozemsko). Relativna pristranskost zaradi nepokritja pri tej spremenljivki je najvišja na Portugalskem, kjer leta 2016 še vedno znaša okrog 10 % (ali 1,64 leta v absolutnem smislu) kljub trendu upadanja. V Sloveniji je relativno pristranskost za to spremenljivko že leta 2010 precej nizka, v obdobju analize pa še upade. Ponovno je treba opozoriti na slabosti te spremenljivke, ki ne meri dejanske stopnje izobrazbe.

Slika 6.14: Relativna pristranskost za politično usmeritev skozi čas



Politična usmeritev je imela med vsemi analiziranimi spremenljivkami najnižjo stopnjo relativne pristranskosti, kar je razvidno tudi iz Slike 6.14. Tu so vrednosti relativne pristranskosti za Slovenijo, vse države in Belgijo, kjer je bila povprečna vrednost najnižja, precej podobne in ne presežejo 2 % relativne pristranskosti. V nekaterih primerih je pristranskost negativna, v drugih pa pozitivna, kar smo opazili tudi že pri spremenljivki spol. Ti rezultati kažejo, da bi za raziskovanje politične usmeritve lahko dokaj brezskrbno uporabili spletne ankete, vendar pa je treba vseeno opozoriti na nekaj faktorjev. Prvi med njimi so posebnosti na nivoju držav. Tako lahko pri Bolgariji opazimo, da so med uporabniki interneta precej bolj zastopani posamezniki z desno politično usmeritvijo, ta pristranskost pa v zadnjih letih še narašča, kljub višji penetraciji interneta. Drugi faktor, ki ga moramo upoštevati je, da se med vsemi izbranimi spremenljivkami prav ta nanaša na najbolj občutljivo temo, na kar kaže že dejstvo, da ima več manjkajočih vrednosti. Medtem ko nekateri anketiranci na vprašanje niso želeli odgovoriti oziroma so odgovorili, da ne vejo, je možno, da so tudi tisti, ki so odgovorili, v večji meri izbrali manj ekstremne vrednosti. Dodatno, v kolikor bi takšno vprašanje postavili v okviru spletne ankete, brez prisotnosti anketarja, lahko pričakujemo še pristranskost zaradi načina anketiranja, saj v primeru spletne ankete pričakujemo manj družbeno zaželenih odgovorov. Pristranskost zaradi nepokritja pri uporabi vzorca le uporabnikov interneta za politično usmerjenost je torej nizka, upoštevati pa moramo še druge vire potencialne pristranskosti.

Slika 6.15: Relativna pristranskost za zadovoljstvo z življenjem skozi čas



Relativna pristranskost za zadovoljstvo z življenjem (gl. Sliko 6.15) ima v našem vzorcu vedno pozitivne vrednosti, kar pomeni, da bi ankete na vzorcu le uporabnikov interneta precenile zadovoljstvo anketirancev. Za vse države je relativna pristranskost nižja od 5 % in z leti rahlo pada, kar pomeni, da pristranskost zaradi nepokritja ne bi bila preveč velika. Za Slovenijo so vrednosti relativne pristranskosti še nižje in ne presegajo 3 % (v absolutnem smislu pristranskost znaša le 0,06 na lestvici od 1 do 4). Država z najnižjimi povprečnimi vrednostmi je ponovno Belgija, ki ima enega največjih deležev penetracije interneta v vzorcu. Bolgarija je ponovno država z najvišjo povprečno relativno pristranskostjo, čeprav v nasprotju s pristranskostjo za politično usmerjenost pri tej spremenljivki pristranskost upada; leta 2016 znaša le še 7 %, kar v absolutnem smislu pomeni pristranskost 0,17 točke na lestvici od 1 do 4.

Rezultati analize relativne pristranskosti za posamezne spremenljivke so pokazali, da večinoma lahko opišemo jasne trende upadanja relativne pristranskosti skozi čas, vendar nekatere države tem trendom ne sledijo, odstopanj pa ne moremo nujno napovedati glede na stopnjo penetracije interneta. Kar se tiče Slovenije, opazamo, da so vrednosti relativne pristranskosti za vse izbrane spremenljivke zelo podobne povprečnim vrednostim za vse analizirane države.

7 Sklep

V magistrskem delu smo se ukvarjali s pristranskostjo zaradi nepokritja v spletnih anketah. Ob pregledu literature smo izpostavili, da sta med glavnimi ovirami za verjetnostne spletne ankete splošne populacije prenizka penetracija interneta in pa razlike med uporabniki in neuporabniki interneta, ki se odražajo v t.i. digitalnem razkoraku. Skupaj prav ta dva faktorja določata, kako visoka bo pristranskost pri specifični anketni oceni, če v anketo ne bi vključili neuporabnikov interneta.

Za potrebe analize pristranskosti smo v magistrskem delu najprej predstavili primere analiz pristranskosti zaradi nepokritja. Te analize uporabljajo metodo simulacije potencialne pristranskosti, saj se izvajajo na podatkih, ki vključujejo tako uporabnike kot tudi neuporabnike interneta. Ti so lahko pridobljeni iz terenskih in telefonskih anket (Couper in drugi 2007; Fuchs in Busse 2009) ali iz verjetnostnih spletnih panelov, ki vključujejo tudi neuporabnike interneta (Leenheer in Scherpenzeel 2013; Eckman 2015; Toepoel in Hendriks 2016). Toepoel in Hendriks (2016) ugotavljata, da v državah z zelo visoko penetracijo interneta, kakršna je Nizozemska, pristranskost zaradi nepokritja ne vpliva pomembno na ocene vrednosti spremenljivk, kljub temu da so dejanske demografske razlike med uporabniki in neuporabniki precej velike. Prav tako analiza Eckman (2015) pokaže, da ob visoki penetraciji interneta pristranskost zaradi nepokritja večinoma ne vpliva na interpretacije rezultatov multivariatnih modelov. Nekaj let prej Leenheer in Scherpenzeel (2013) ugotavljata, da so verjetnostni spletni paneli bolj reprezentativni od anket, ki bi izključile uporabnike interneta, vendar opozarjata tudi na njihove druge slabosti, kot so npr. pristranskost zaradi nizke začetne stopnje odgovora, ki je pri rekrutiranju udeležencev veliko nižja za gospodinjstva brez interneta.

Pregled penetracije interneta na podlagi podatkov iz Evrobarometra je pokazal, da je delež uporabnikov interneta od leta 2010 do leta 2016 narasel za 12 odstotnih točk. Večji del te rasti se je zgodil med leti 2010 in 2013, nato pa se je upočasnila, ponekod pa je delež anketirancev, ki imajo doma internetno povezavo, celo upadel. Delno lahko nihanja v penetraciji interneta v naših podatkih pripišemo anketni napaki, delno pa gre lahko tudi za druge faktorje, kot je čedalje večja razširjenost mobilnega interneta, ki ni eksplicitno vključena v Evrobarometrovem vprašanju o dostopu do interneta. Pri primerjavi z rezultati v delu Mohorko in drugih (2011), po katerem smo se zgledovali v empiričnem delu

magistrskega dela, se je izkazalo, da je bila rast penetracije interneta v obdobju med leti 2005 in 2009 hitrejša in nekoliko bolj konsistentna.

Z bivariatnimi analizami smo pokazali, da v letih, ki jih pokriva naša analiza, ni prišlo do pomembnih sprememb v značilnostih uporabnikov in neuporabnikov interneta. Uporabniki interneta so še vedno v povprečju mlajši od neuporabnikov in so redno izobraževanje končali pri višji starosti. Med njimi je večji delež moških. Razlike v zadovoljstvu z življenjem so bile precej konsistentne iz leta v leto; uporabniki interneta so nekoliko bolj zadovoljni od neuporabnikov. Slovenija odstopa od celotnega vzorca pri povprečni starosti neuporabnikov interneta, ki je kar precej višja. Med preučevanimi spremenljivkami smo najmanjše razlike med uporabniki in neuporabniki našli pri politični usmeritvi: v Sloveniji so se izkazale za značilne samo v letu 2016, ko se je pokazalo, da so uporabniki interneta politično usmerjeni rahlo bolj levo od neuporabnikov.

Glavno raziskovalno vprašanje tega magistrskega dela je bilo, kako bi pristranskost zaradi nepokritja vplivala na spletne ankete ali spletne (neverjetnostne) panele splošne populacije, iz katerih bi izključili neuporabnike interneta. Analiza pristranskosti na podatkih za države EU, Turčijo in turško skupnost na Cipru je pokazala, da bi v takem primeru podcenili povprečno starost v splošni populaciji ter precenili povprečno starost pri končani izobrazbi in zadovoljstvo z življenjem. Medtem ko je bila pristranskost največja pri spremenljivki starost, je bila najmanjša pri spremenljivkah spol in politična usmeritev, kjer bi le rahlo podcenili delež žensk v splošni populaciji in nagnjenje k politični desnici. Dejanski učinek na anketne ocene si najlažje predstavljamo, če si ogledamo vrednosti absolutne pristranskosti. Ta v povprečju znaša -1,5 % za spol, -4,3 leta za starost, 1,1 leta za starost pri dokončani izobrazbi, -0,1 na lestvici od 1 (najbolj levo) do 10 (najbolj desno) za politično usmeritev in 0,1 na lestvici od 1 (sploh ni zadovoljen) do 4 (zelo zadovoljen) za zadovoljstvo z življenjem. Enake ugotovitve ne veljajo nujno na nivoju posameznih držav. Čeprav so bili rezultati analize pristranskosti za Slovenijo precej podobni povprečju za vse države, smo pri drugih državah opazili večja odstopanja, ki niso bila odvisna le od stopnje penetracije interneta. Primer je Bolgarija, kjer najdemo visoko pristranskost zaradi nepokritja pri spremenljivki politična usmeritev, čeprav je to spremenljivka z najmanjšo pristranskostjo na nivoju vzorca vseh držav.

Zanimalo nas je tudi, kako se je pristranskost zaradi nepokritja spremenila od leta 2010 do leta 2016 in kako se to obdobje razlikuje od obdobja med leti 2005 in 2009, ki ga pokrivajo Mohorko in drugi (2011). Pri starosti, starosti pri končani izobrazbi in zadovoljstvu z življenjem, kjer je nivo pristranskosti na splošno višji, lahko opazimo upadanje pristranskosti skozi leta, ki je tesno povezano z naraščanjem penetracije interneta. Pri spremenljivkah spol in politična usmeritev, kjer je nivo pristranskosti že tako nizek, namesto upadanja opazimo ohranjanje podobnega povprečnega nivoja pristranskosti z nekaj nihanjem. Povečini rezultati sledijo trendom, ki so jih opisali že Mohorko in drugi (2011), vendar pa je v drugem obdobju med leti 2010 in 2016 upadanje pristranskosti počasnejše kot v prvem obdobju med leti 2005 in 2009, kar sovпада z upočasnitvijo rasti penetracije interneta. Zadnji podatki iz leta 2016 kažejo, da absolutna pristranskost pri spremenljivki spol znaša -1,1 %, pri spremenljivki starost -4,1 leta, pri spremenljivki starost pri dokončani izobrazbi 0,9 leta, pri spremenljivki politična usmeritev le -0,04 točke na lestvici od 1 do 10, pri spremenljivki zadovoljstvo z življenjem pa 0,1 točke na lestvici od 1 do 4.

S praktičnega vidika so rezultati takšnih analiz zanimivi zaradi vprašanja učinkovitosti finančnih vložkov v tehnično opremo, s katero bi neuporabnikom interneta omogočili sodelovanje v verjetnostnih spletnih panelih. Že pregled literature je pokazal, da takšne investicije v državah z zelo visokim nivojem penetracije interneta (npr. na Nizozemskem, kjer nivo penetracije interneta že kar nekaj časa presega 90 %) nimajo več pomembnega učinka na pridobljene podatke, saj je pristranskost zaradi nepokritja že sama po sebi majhna. Kljub rasti penetracije interneta v Evropi je takšnih držav še vedno le peščica, kar pomeni, da je treba pri zasnovi anket še vedno upoštevati pristranskost zaradi nepokritja. To je še zlasti pomembno pri študijah, ki se ukvarjajo s spremenljivkami, kjer gre pričakovati višjo pristranskost (npr. s starostjo). Podatki tudi kažejo, da se digitalni razkorak lahko zapre samo do neke meje, majhen odstotek pristranskosti zaradi razlik med uporabniki in neuporabniki interneta pa se bo verjetno ohranil. V prihodnosti bo verjetno vse bolj pomembno upoštevati tudi sekundarne oblike digitalnega razkoraka; tudi če bomo imeli praktično vsi dostop do interneta, ta verjetno ne bo enak za vse.

Z vidika analiz pristranskosti zaradi nepokritja v spletnih anketah bi bilo v prihodnje priporočljivo bolj upoštevati tudi mobilne širokopasovne povezave do interneta. Že Fuchs in Busse (2009) namreč ugotavljata, da je ponekod v Evropi pristranskost zaradi nepokritja manjša, če vključimo le uporabnike mobilnega interneta, kot če vključimo le uporabnike

stacionarnega interneta. V Evrobarometrih anketah se od leta 2014 poleg spremenljivke *anketiranec ima doma internetno povezavo* pojavlja spremenljivka *anketiranec ima pametni telefon*, ki bi jo bilo smiselno vključiti v nadaljnje analize kot indikator dostopa do mobilnega interneta. V tem magistrskem delu se za ta korak nismo odločili, ker smo želeli ohraniti čim večjo primerljivost z metodologijo Mohorko in drugih (2011). Ta metodološka omejitev se odraža v podatkih, ki kažejo upadanje penetracije interneta v nekaterih državah (najbolj izrazito v Italiji), in prav gotovo predstavlja eno od glavnih pomanjkljivosti naše analize. Upoštevanje hitrega razvoja na področju informacijsko-komunikacijskih tehnologij tako ostaja velik izziv za raziskovanje sprememb pristranskosti zaradi nepokritja v spletnih anketah v nekem časovnem obdobju.

8 Literatura

- Bethlehem, J. in S. Biffignandi. 2012. *Handbook of Web Surveys*. Hoboken, New Jersey: John Wiley & Sons, Inc.
- Biemer, Paul in Lars Lyberg. 2003. *Introduction to survey quality*. Hoboken, New Jersey: John Wiley & Sons, Inc.
- Callegaro, Mario, Katja Lozar Manfreda in Vasja Vehovar. 2015. *Web Survey Methodology*. London: Sage.
- Couper, Mick, Arie Kapteyn, Matthias Schonlau in Joachim Winter. 2007. Noncoverage and Nonresponse in an Internet Survey. *Social Science Research* 36: 131–148.
- Couper, Mick in Peter Miller. 2008. Web Survey Methods: Introduction. *Public Opinion Quarterly* 72 (5): 831–35.
- Couper, Mick. 2000. Web Surveys: A Review of Issues and Approaches. *Public Opinion Quarterly* 64: 464–494.
- de Leeuw, Edith, Joop Hox in Don Dillman. 2009. *International handbook of survey methodology*. New York in London: Psychology Press.
- de Leeuw, Edith. 2012. Counting and Measuring Online: The Quality of Internet Surveys. *Bulletin of Sociological Methodology* 114 (1): 68–78.
- Dolničar, Vesna, Katja Vukčević, Luka Kronegger in Vasja Vehovar. 2002. Digitalni razkorak v Sloveniji. *Družboslovne razprave* 18 (40): 83–106.
- Eckman, Stephanie. 2016. Does the Inclusion of Non-Internet Households in a Web Panel Reduce Coverage Bias? *Social Science Computer Review* 34 (1): 41–58.
- Eurostat. 2017a. *Households - level of internet access*. Dostopno prek: http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=isoc_ci_in_h&lang=en (31. julij 2017).
- --- 2017b. *Individuals - internet use*. Dostopno prek: http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=isoc_ci_ifp_iu&lang=en (31. julij 2017).
- --- 2017c. *ICT usage in households and by individuals: Reference metadata*. Dostopno prek: http://ec.europa.eu/eurostat/cache/metadata/en/isoc_i_esms.htm (16. avgust 2017).

- Evropska komisija in Evropski parlament. 2012. *Eurobarometer 73.1 (Jan-Feb 2010)*. Datoteka podatkov. TNS opinion, Bruselj. Köln: GESIS Data Archive. Dostopno prek: doi:10.4232/1.11428 (11. avgust 2017).
- Evropska komisija. 2012. *Eurobarometer 73.5 (Jun 2010)*. Datoteka podatkov. TNS opinion, Bruselj. Köln: GESIS Data Archive. Dostopno prek: doi:10.4232/1.11432 (11. avgust 2017).
- --- 2014. *Eurobarometer 75.3 (2011)*. Datoteka podatkov. TNS opinion, Bruselj. Köln: GESIS Data Archive. Dostopno prek: doi:10.4232/1.11852 (11. avgust 2017).
- --- 2015a. *Eurobarometer 77.3 (2012)*. Datoteka podatkov. TNS opinion, Bruselj. Köln: GESIS Data Archive. Dostopno prek: doi:10.4232/1.12050 (11. avgust 2017).
- --- 2015b. *Eurobarometer 77.4 (2012)*. Datoteka podatkov. TNS opinion, Bruselj. Köln: GESIS Data Archive. Dostopno prek: doi: 10.4232/1.12049 (11. avgust 2017).
- --- 2015c. *Eurobarometer 82.3 (2014)*. Datoteka podatkov. TNS opinion, Bruselj. Köln: GESIS Data Archive. Dostopno prek: doi:10.4232/1.12259 (11. avgust 2017).
- --- 2017a. *Eurobarometer 79.5 (2013)*. Datoteka podatkov. TNS opinion, Bruselj. Köln: GESIS Data Archive. Dostopno prek: doi:10.4232/1.12738 (11. avgust 2017).
- --- 2017b. *Eurobarometer 80.1 (2013)*. Datoteka podatkov. TNS opinion, Bruselj. Köln: GESIS Data Archive. Dostopno prek: doi:10.4232/1.12768 (11. avgust 2017).
- --- 2017c. *Eurobarometer 84.3 (2015)*. Datoteka podatkov. TNS opinion, Bruselj. Köln: GESIS Data Archive. Dostopno prek: doi:10.4232/1.12799 (11. avgust 2017).
- --- 2017č. *Eurobarometer 86.2 (2016)*. Datoteka podatkov. TNS opinion, Bruselj. Köln: GESIS Data Archive. Dostopno prek: doi:10.4232/1.12801 (11. avgust 2017).
- Fuchs, Marek in Britta Busse. 2009. The Coverage Bias of Mobile Web Surveys Across European Countries. *International Journal of Internet Science* 4 (1): 21–33.

- Groves, Robert M., Floyd J. Fowler, Jr., Mick P. Couper, James M. Lepkowski, Eleanor Singer, Roger Tourangeau. 2009. *Survey methodology*. Hoboken: Wiley.
- Leenheer, J in A. C. Scherpenzeel. 2013. Does It Pay Off to Include Non-Internet Households in an Internet Panel? *International Journal of Internet Science* 8 (1): 17–29.
- *LISS Panel Data*. 2009. Dostopno prek: <https://www.lissdata.nl/> (27. julij 2017).
- Lozar Manfreda, Katja in Vasja Vehovar. 2009. Internet Surveys. V *International handbook of survey methodology*, ur. Edith de Leeuw, Edith, Joop Hox in Don Dillman, 264–284. New York in London: Psychology Press.
- Lynn, Peter. 2009. *Methodology of longitudinal surveys*. Chichester: Wiley.
- Mohorko, Anja, Edith de Leeuw in Joop Hox. 2011. *Internet Coverage and Coverage Bias in Countries Across Europe and over Time: Background, Methods, Question Wording and Bias Tables*. Dostopno prek: <http://www.joopbox.net/papers/WebCoverage.pdf> (15. maj 2017).
- Toepoel, Vera in Yannie Hendriks. 2016. The Impact of Non-Coverage in Web Surveys in a Country with High Internet Penetration: Is It (Still) Useful to Provide Equipment to Non-Internet Households in the Netherlands? *International Journal of Internet Science* 11 (1): 33–50.
- Villar, Ana. 2016. A survey future online: Establishing the Cross-National Online Survey Panel. *SERISS Newsletter* 1. Dostopno prek: <https://seriss.eu/news/newsletter/survey-future-online-establishing-cronos/> (16. avgust 2017).
- *Weighting overview, Eurobarometer Data Service*. Dostopno prek: <https://www.gesis.org/eurobarometer-data-service/survey-series/standard-special-eb/weighting-overview/> (8. september 2017).

Priloge

Priloga A: Uporabljena anketna vprašanja in iz njih izpeljane spremenljivke

Spremenljivka	Oznaka	Vprašanje	Originalne vrednosti	Vrednosti po rekodiranju
Spol	D10	Spol.	1-moški 2-ženski	0-moški 1-ženski
Starost	D11	Koliko ste stari?		/
Starost pri dokončani izobrazbi	D8	Koliko ste bili stari, ko ste zaključili redno izobraževanje?		/
Politična usmeritev	D1	Pri političnih zadevah ljudje govorijo o »levici« in »desnici«. Kam bi vi na tej lestvici umestili svoje poglede?	1-levo 2 3 4 5 6 7 8 9 10-desno	/
Zadovoljstvo z življenjem	D70 ali QA1 (leta 2010, 2011 in 2013) ali QB1 (leta 2012)	Gledano v celoti, ali ste zelo zadovoljni, razmeroma zadovoljni, ne preveč zadovoljni ali sploh niste zadovoljni s svojim življenjem?	1- Zelo zadovoljni 2- Razmeroma zadovoljni 3- Ne preveč zadovoljni 4- Sploh niste zadovoljni	1- Sploh niste zadovoljni 2- Ne preveč zadovoljni 3- Razmeroma zadovoljni 4- Zelo zadovoljni
Doma ima internetno povezavo	D46	Katere izmed sledečih potrošnih dobrin imate? Internetno povezavo doma	0-Je omenjeno 1-Ni omenjeno	/

Priloga B: Penetracija interneta v gospodinjstvih za obdobje 2010–2016

	2010	2011	2012	2013	2014	2015	2016
Avstrija	63,3 %	68,1 %	71,0 %	76,6 %	73,1 %	72,8 %	76,6 %
Belgija	71,7 %	73,2 %	75,3 %	81,2 %	81,6 %	82,3 %	81,6 %
Bolgarija	43,1 %	55,4 %	59,7 %	62,6 %	51,6 %	59,8 %	55,8 %
Hrvaška	54,8 %	57,5 %	59,5 %	65,5 %	67,3 %	63,3 %	64,1 %
TCC		57,8 %	69,0 %	77,8 %	73,2 %	77,6 %	72,6 %
Republika Ciper	61,2 %	61,9 %	64,2 %	77,7 %	71,0 %	72,6 %	71,6 %
Češka	59,8 %	60,9 %	66,5 %	71,1 %	70,3 %	70,4 %	73,3 %
Danska	85,9 %	88,9 %	92,9 %	94,1 %	94,1 %	94,4 %	92,5 %
Estonija	68,9 %	72,0 %	74,2 %	78,4 %	77,9 %	75,6 %	76,3 %
Finska	81,5 %	84,3 %	83,8 %	85,8 %	84,0 %	84,9 %	84,5 %
Francija	68,6 %	74,2 %	78,2 %	80,9 %	79,2 %	80,8 %	82,4 %
Nemčija - Vzhod	65,0 %	67,5 %	65,9 %	66,3 %	69,8 %	71,6 %	71,7 %
Nemčija - Zahod	71,2 %	75,3 %	78,8 %	81,3 %	76,8 %	81,1 %	75,9 %
Velika Britanija	74,3 %	76,3 %	78,6 %	80,6 %	79,3 %	83,9 %	82,0 %
Grčija	40,6 %	47,7 %	54,7 %	61,4 %	64,5 %	64,8 %	70,1 %
Madžarska	46,7 %	46,5 %	54,2 %	57,7 %	57,3 %	60,3 %	64,2 %
Irska	74,9 %	78,9 %	81,9 %	76,6 %	78,3 %	82,9 %	83,9 %
Italija	53,6 %	60,0 %	60,3 %	64,5 %	53,6 %	49,3 %	52,4 %
Latvija	53,0 %	63,2 %	69,0 %	75,7 %	70,4 %	72,1 %	75,2 %
Litva	51,4 %	54,9 %	59,8 %	62,7 %	67,5 %	69,1 %	66,4 %
Luksemburg	83,9 %	84,6 %	84,4 %	83,5 %	84,4 %	83,6 %	89,6 %
Malta	74,8 %	75,4 %	81,0 %	80,4 %	80,6 %	84,9 %	82,9 %
Severna Irska	70,2 %	64,7 %	73,7 %	74,9 %	73,9 %	84,7 %	79,7 %
Poljska	61,5 %	61,3 %	61,7 %	68,2 %	68,1 %	64,8 %	65,3 %
Portugalska	44,0 %	51,2 %	44,6 %	49,8 %	57,2 %	63,8 %	66,2 %
Romunija	39,2 %	37,7 %	47,2 %	53,7 %	54,2 %	53,0 %	54,2 %
Slovaška	54,1 %	63,9 %	68,3 %	73,4 %	67,1 %	71,0 %	70,1 %
Slovenija	71,3 %	71,5 %	74,3 %	78,6 %	75,5 %	75,8 %	78,0 %
Španija	52,8 %	58,5 %	60,1 %	68,1 %	70,1 %	66,8 %	68,5 %
Švedska	90,4 %	91,4 %	93,3 %	95,6 %	90,8 %	93,8 %	90,8 %
Nizozemska	93,1 %	92,5 %	94,7 %	95,2 %	96,8 %	97,4 %	98,1 %
Turčija	34,2 %	38,9 %	37,7 %	51,6 %	54,0 %	51,9 %	54,8 %

Priloga C: Relativna in absolutna pristranskost zaradi nepokritja po letih in državah

Tabela C.1: Relativna pristranskost zaradi nepokritja za spremenljivko *spol*

Lestvica: 0 (Moški), 1 (Ženski)

Negativni predznak: vzorčni okvir izključno uporabnikov interneta precenjuje delež moških.

Država/Leto	2010	2011	2012	2013	2014	2015	2016	Povprečje
Avstrija	-0,081	-0,079	-0,044	-0,036	-0,032	-0,025	-0,022	-0,045
Belgija	-0,042	-0,023	-0,018	-0,018	-0,055	-0,015	-0,019	-0,027
Bolgarija	0,002	-0,056	-0,015	0,017	-0,054	-0,025	-0,035	-0,024
Hrvaška	0,006	-0,005	-0,007	-0,005	-0,044	-0,029	-0,022	-0,015
Republika Ciper	-0,014	-0,021	-0,001	-0,034	-0,032	-0,040	-0,034	-0,025
TCC		-0,007	-0,059	-0,004	-0,014	-0,065	-0,015	-0,027
Češka	-0,066	-0,067	-0,048	-0,065	-0,074	-0,037	-0,010	-0,052
Danska	-0,026	-0,011	-0,019	-0,011	-0,014	-0,007	-0,012	-0,014
Estonija	-0,082	-0,037	-0,044	-0,021	-0,023	-0,032	-0,068	-0,044
Finska	-0,040	-0,031	-0,044	-0,035	-0,027	-0,045	-0,016	-0,034
Francija	-0,062	-0,036	-0,045	-0,023	-0,025	0,004	-0,008	-0,028
Nemčija – Vzhod	-0,081	-0,088	-0,037	-0,024	-0,014	-0,077	-0,034	-0,051
Nemčija – Zahod	-0,047	-0,051	-0,039	-0,043	-0,034	-0,041	-0,043	-0,043
Velika Britanija	-0,040	-0,048	-0,011	-0,031	-0,035	-0,022	-0,013	-0,028
Grčija	-0,093	-0,044	-0,070	-0,099	-0,049	-0,061	-0,014	-0,061
Madžarska	-0,037	-0,048	-0,043	0,015	-0,018	-0,056	-0,051	-0,034
Irska	-0,002	-0,007	0,004	0,032	-0,011	0,008	-0,002	0,003
Italija	-0,055	0,002	0,025	-0,029	-0,104	-0,067	-0,018	-0,035
Latvija	-0,006	-0,032	-0,012	0,006	-0,014	-0,030	-0,036	-0,018
Litva	-0,031	0,009	-0,054	-0,045	-0,040	-0,043	-0,046	-0,036
Luksemburg	-0,063	-0,053	-0,040	-0,011	-0,013	-0,062	-0,009	-0,036
Malta	0,015	-0,037	-0,031	-0,052	-0,055	-0,046	-0,015	-0,032
Nizozemska	-0,014	-0,001	-0,004	-0,005	-0,009	-0,005	0,002	-0,005
Severna Irska	-0,057	0,019	0,053	-0,095	0,005	-0,042	-0,008	-0,018
Poljska	-0,029	-0,031	-0,015	-0,029	-0,038	-0,028	-0,030	-0,029
Portugalska	-0,054	-0,041	0,002	-0,020	-0,041	-0,027	-0,024	-0,029
Romunija	-0,040	-0,015	-0,007	-0,012	-0,050	-0,066	-0,027	-0,031
Slovaška	-0,091	-0,043	-0,017	-0,053	-0,026	-0,048	-0,040	-0,045
Slovenija	-0,043	-0,043	-0,021	-0,041	-0,039	-0,049	-0,038	-0,039
Španija	-0,041	-0,035	-0,039	-0,026	-0,087	-0,049	-0,027	-0,043
Švedska	-0,024	-0,024	-0,007	-0,010	-0,035	-0,020	-0,042	-0,023
Turčija	-0,072	-0,056	-0,067	0,058	-0,074	0,047	-0,039	-0,029
EU27	-0,047	-0,036	-0,023	-0,029	-0,045	-0,033	-0,024	-0,034
EU28+TUR+TCC		-0,035	-0,024	-0,020	-0,046	-0,025	-0,022	-0,029

Tabela C.2: Relativna pristranskost zaradi nepokritja za spremenljivko starost

Negativni predznak: vzorčni okvir izključno uporabnikov interneta podcenjuje povprečno starost

Država/Leto	2010	2011	2012	2013	2014	2015	2016	Povprečje
Avstrija	-0,132	-0,138	-0,130	-0,118	-0,120	-0,126	-0,097	-0,123
Belgija	-0,098	-0,100	-0,076	-0,071	-0,083	-0,081	-0,077	-0,084
Bolgarija	-0,198	-0,192	-0,167	-0,165	-0,196	-0,179	-0,178	-0,182
Hrvaška	-0,154	-0,150	-0,154	-0,152	-0,136	-0,108	-0,097	-0,136
Republika Ciper	-0,180	-0,170	-0,138	-0,113	-0,114	-0,138	-0,119	-0,139
TCC		-0,124	-0,095	-0,106	-0,078	-0,106	-0,035	-0,091
Češka	-0,154	-0,149	-0,127	-0,113	-0,117	-0,118	-0,098	-0,125
Danska	-0,064	-0,056	-0,036	-0,036	-0,031	-0,025	-0,044	-0,042
Estonija	-0,136	-0,165	-0,128	-0,122	-0,103	-0,108	-0,119	-0,126
Finska	-0,087	-0,074	-0,086	-0,065	-0,073	-0,054	-0,060	-0,071
Francija	-0,113	-0,095	-0,077	-0,082	-0,083	-0,084	-0,065	-0,085
Nemčija - Vzhod	-0,127	-0,126	-0,119	-0,135	-0,104	-0,122	-0,083	-0,117
Nemčija - Zahod	-0,125	-0,108	-0,083	-0,083	-0,094	-0,077	-0,073	-0,092
Velika Britanija	-0,090	-0,078	-0,086	-0,065	-0,044	-0,066	-0,060	-0,070
Grčija	-0,205	-0,184	-0,161	-0,187	-0,143	-0,157	-0,137	-0,168
Madžarska	-0,153	-0,139	-0,135	-0,123	-0,124	-0,136	-0,122	-0,133
Irska	-0,082	-0,069	-0,074	-0,073	-0,064	-0,060	-0,058	-0,069
Italija	-0,091	-0,084	-0,105	-0,109	-0,104	-0,134	-0,128	-0,108
Latvija	-0,119	-0,122	-0,103	-0,092	-0,106	-0,105	-0,088	-0,105
Litva	-0,204	-0,195	-0,176	-0,164	-0,155	-0,149	-0,145	-0,170
Luksemburg	-0,073	-0,066	-0,064	-0,063	-0,030	-0,051	-0,040	-0,055
Malta	-0,106	-0,121	-0,107	-0,094	-0,094	-0,084	-0,081	-0,098
Nizozemska	-0,031	-0,019	-0,026	-0,020	-0,016	-0,012	-0,007	-0,019
Severna Irska	-0,099	-0,090	-0,096	-0,086	-0,062	-0,089	-0,050	-0,082
Poljska	-0,155	-0,161	-0,144	-0,149	-0,125	-0,134	-0,138	-0,144
Portugalska	-0,164	-0,201	-0,199	-0,200	-0,174	-0,145	-0,155	-0,177
Romunija	-0,195	-0,172	-0,180	-0,164	-0,154	-0,170	-0,172	-0,172
Slovaška	-0,162	-0,133	-0,123	-0,119	-0,119	-0,113	-0,122	-0,127
Slovenija	-0,135	-0,136	-0,126	-0,109	-0,100	-0,112	-0,080	-0,114
Španija	-0,175	-0,144	-0,146	-0,128	-0,113	-0,133	-0,115	-0,136
Švedska	-0,049	-0,043	-0,034	-0,023	-0,037	-0,018	-0,037	-0,034
Turčija	-0,111	-0,115	-0,129	-0,133	-0,106	-0,125	-0,057	-0,111
EU27	-0,116	-0,105	-0,100	-0,097	-0,092	-0,098	-0,090	-0,100
EU28+TUR+TCC		-0,099	-0,094	-0,095	-0,089	-0,095	-0,086	-0,093

Tabela C.3: Relativna pristranskost zaradi nepokritja za spremenljivko starost pri končani izobrazbi

Pozitivni predznak: vzorčni okvir izključno uporabnikov interneta precenjuje povprečno starost pri končani izobrazbi

Država/Leto	2010	2011	2012	2013	2014	2015	2016	Povprečje
Avstrija	0,033	0,037	0,034	0,025	0,041	0,028	0,032	0,033
Belgija	0,042	0,039	0,037	0,035	0,021	0,021	0,034	0,033
Bolgarija	0,077	0,056	0,071	0,058	0,072	0,052	0,059	0,063
Hrvaška	0,062	0,048	0,052	0,038	0,031	0,022	0,032	0,041
Republika Ciper	0,112	0,100	0,095	0,051	0,080	0,075	0,063	0,082
TCC		0,056	0,056	0,057	0,037	0,054	0,012	0,045
Češka	0,027	0,025	0,026	0,023	0,022	0,025	0,020	0,024
Danska	0,022	0,017	0,012	0,013	0,006	0,000	0,006	0,011
Estonija	0,033	0,024	0,021	0,022	0,026	0,025	0,029	0,026
Finska	0,036	0,033	0,030	0,023	0,033	0,024	0,040	0,031
Francija	0,051	0,044	0,040	0,038	0,035	0,033	0,020	0,037
Nemčija - Vzhod	0,052	0,041	0,035	0,040	0,038	0,048	0,040	0,042
Nemčija - Zahod	0,045	0,032	0,028	0,036	0,034	0,027	0,035	0,034
Velika Britanija	0,031	0,030	0,030	0,027	0,024	0,021	0,019	0,026
Grčija	0,152	0,133	0,113	0,096	0,103	0,104	0,091	0,113
Madžarska	0,076	0,072	0,065	0,056	0,060	0,043	0,047	0,060
Irska	0,037	0,028	0,034	0,037	0,026	0,020	0,024	0,029
Italija	0,084	0,081	0,083	0,086	0,086	0,107	0,092	0,089
Latvija	0,026	0,033	0,022	0,018	0,027	0,020	0,022	0,024
Litva	0,057	0,056	0,040	0,050	0,043	0,036	0,048	0,047
Luksemburg	0,029	0,027	0,021	0,020	0,022	0,027	0,022	0,024
Malta	0,033	0,038	0,035	0,035	0,034	0,036	0,025	0,034
Nizozemska	0,010	0,009	0,007	0,007	0,008	0,007	0,004	0,007
Severna Irska	0,037	0,038	0,035	0,033	0,019	0,026	0,026	0,031
Poljska	0,052	0,060	0,058	0,047	0,046	0,040	0,042	0,049
Portugalska	0,168	0,130	0,147	0,136	0,102	0,104	0,100	0,127
Romunija	0,111	0,101	0,094	0,068	0,086	0,057	0,064	0,083
Slovaška	0,040	0,030	0,028	0,025	0,031	0,027	0,024	0,029
Slovenija	0,048	0,046	0,038	0,032	0,027	0,028	0,031	0,036
Španija	0,104	0,099	0,088	0,076	0,065	0,075	0,076	0,083
Švedska	0,023	0,013	0,011	0,007	0,016	0,000	0,001	0,010
Turčija	0,080	0,109	0,093	0,086	0,085	0,088	0,074	0,088
EU27	0,064	0,058	0,055	0,052	0,049	0,047	0,046	0,053
EU28+TUR+TCC		0,068	0,067	0,060	0,055	0,055	0,046	0,059

Tabela C.4: Relativna pristranskost zaradi nepokritja za spremenljivko *politična usmeritev levo-desno*

Lestvica: 1 (levo) – 10 (desno)

Negativni predznak: vzorčni okvir izključno uporabnikov interneta precenjuje levo politično usmerjenost

Država/Leto	2010	2011	2012	2013	2014	2015	2016	Povprečje
Avstrija	-0,005	0,015	-0,002	-0,014	-0,016	-0,016	-0,029	-0,009
Belgija	-0,005	-0,017	0,002	0,007	0,005	0,008	0,004	0,001
Bolgarija	0,103	0,099	0,108	0,113	0,147	0,164	0,188	0,132
Hrvaška	0,000	-0,013		-0,011	-0,006	-0,044	-0,029	-0,017
Republika Ciper	0,031	0,030	-0,037	0,026	0,007	0,007	-0,008	0,008
TCC		-0,067			-0,023	-0,040	-0,032	-0,040
Češka	0,079	0,095	0,076	0,061	0,064	0,070	0,056	0,072
Danska	-0,018	0,001	-0,005	-0,004	0,000	-0,006	-0,007	-0,006
Estonija	0,017	0,017	0,006	0,018	0,005	0,019	0,011	0,013
Finska	0,002	0,004	0,005	0,006	0,006	0,005	-0,003	0,004
Francija	-0,017	-0,009	0,001	-0,016	-0,014	-0,004	-0,018	-0,011
Nemčija - Vzhod	-0,002	-0,021	0,018	-0,019	0,018	0,009	-0,009	-0,001
Nemčija - Zahod	-0,029	-0,025	-0,011	-0,020	-0,018	-0,013	-0,018	-0,019
Velika Britanija	0,003	0,001	-0,009	-0,013	-0,001	-0,010	0,002	-0,004
Grčija	-0,055	-0,043	-0,041	-0,028	-0,020	-0,044	-0,008	-0,034
Madžarska	0,036	0,039	0,045	0,016	0,027	0,060	0,017	0,034
Irska	-0,011	-0,022	-0,012	0,022	-0,014	-0,011	0,006	-0,006
Italija	0,011	0,026	0,015	0,015	0,018	-0,057	-0,014	0,002
Latvija	0,017	0,005	0,004	-0,008	-0,001	0,008	0,009	0,005
Litva	-0,025	0,011	-0,010	0,021	-0,009	-0,002	0,035	0,003
Luksemburg	-0,024	-0,019	-0,008	-0,003	-0,012	-0,013	-0,002	-0,011
Malta	-0,021	0,024	0,015	-0,008	0,022	0,005	0,011	0,007
Nizozemska	-0,005	-0,004	0,000	-0,011	0,001	-0,003	-0,001	-0,003
Severna Irska	-0,007	-0,027	-0,015	-0,007	-0,023	0,003	-0,004	-0,011
Poljska	-0,008	-0,002	-0,007	0,002	-0,013	-0,036	-0,014	-0,011
Portugalska	-0,022	0,016	-0,004	0,030	0,026	0,014	0,025	0,012
Romunija	0,035	0,039	0,040	0,021	0,045	0,079	0,046	0,044
Slovaška	0,033	0,046	0,046	0,038	0,044	0,034	0,072	0,045
Slovenija	0,004	0,003	-0,009	-0,008	-0,015	-0,013	-0,023	-0,009
Španija	-0,018	-0,003	-0,052	-0,007	-0,023	-0,048	-0,023	-0,025
Švedska	0,003	-0,001	0,000	-0,004	0,009	0,007	0,004	0,003
Turčija	-0,067	-0,045			-0,018	-0,036	-0,025	-0,038
EU27	-0,008	0,000	-0,001	-0,003	-0,003	-0,012	-0,005	-0,005
EU28+TUR+TCC		-0,017			-0,007	-0,021	-0,007	-0,013

Tabela C.5: Relativna pristranskost zaradi nepokritja za spremenljivko zadovoljstvo z življenjem

Lestvica: 1 (sploh ni zadovoljen/zadovoljna) – 4 (zelo zadovoljen/zadovoljna)

Pozitivni predznak: vzorčni okvir izključno uporabnikov interneta precenjuje povprečno zadovoljstvo

Država/Leto	2010	2011	2012	2013	2014	2015	2016	Povprečje
Avstrija	0,024	0,021	0,012	0,017	-0,009	0,012	0,018	0,014
Belgija	0,019	0,013	0,020	0,014	0,009	0,017	0,006	0,014
Bolgarija	0,138	0,106	0,093	0,092	0,086	0,073	0,070	0,094
Hrvaška		0,069	0,058	0,050	0,039	0,029	0,018	0,044
Republika Ciper	0,021	0,051	0,040	0,019	0,034	0,026	0,025	0,031
TCC		0,028	0,021	0,028	0,013	0,010	0,010	0,019
Češka	0,042	0,037	0,024	0,030	0,024	0,026	0,019	0,029
Danska	0,007	0,007	0,004	0,006	0,001	0,001	0,004	0,004
Estonija	0,020	0,032	0,033	0,025	0,028	0,024	0,029	0,027
Finska	0,011	0,005	0,008	0,009	0,006	0,003	0,007	0,007
Francija	0,025	0,013	0,021	0,020	0,021	0,015	0,015	0,018
Nemčija - Vzhod	0,017	0,025	0,015	0,028	0,019	0,015	0,015	0,019
Nemčija - Zahod	0,022	0,007	0,017	0,009	0,009	0,006	0,015	0,012
Velika Britanija	0,009	0,017	0,008	0,007	0,008	0,008	0,009	0,009
Grčija	0,073	0,103	0,068	0,053	0,063	0,047	0,042	0,064
Madžarska	0,062	0,071	0,048	0,047	0,058	0,034	0,041	0,051
Irska	0,011	0,007	0,008	0,018	0,007	0,010	0,007	0,010
Italija	0,041	0,028	0,029	0,018	0,026	0,038	0,016	0,028
Latvija	0,050	0,060	0,034	0,032	0,033	0,028	0,030	0,038
Litva	0,072	0,066	0,047	0,042	0,046	0,039	0,035	0,050
Luksemburg	0,003	0,005	-0,004	0,006	-0,004	0,011	0,001	0,002
Malta	0,006	0,030	0,014	0,022	0,004	0,006	0,004	0,012
Nizozemska	0,007	0,005	0,003	0,000	0,002	0,002	0,001	0,003
Severna Irska	0,005	0,015	0,013	-0,008	0,010	-0,002	0,008	0,006
Poljska	0,042	0,037	0,035	0,039	0,028	0,026	0,037	0,035
Portugalska	0,078	0,069	0,073	0,070	0,068	0,025	0,044	0,061
Romunija	0,119	0,099	0,087	0,085	0,054	0,064	0,071	0,083
Slovaška	0,043	0,035	0,055	0,035	0,037	0,026	0,024	0,037
Slovenija	0,022	0,033	0,016	0,018	0,018	0,017	0,024	0,021
Španija	0,028	0,030	0,039	0,039	0,038	0,024	0,032	0,033
Švedska	0,009	0,003	0,001	0,003	0,002	0,004	0,001	0,003
Turčija		0,019	0,021	-0,003	0,016	-0,009	-0,008	0,006
EU27	0,047	0,042	0,043	0,038	0,034	0,034	0,032	0,038
EU28+TUR+TCC		0,045	0,043	0,036	0,035	0,033	0,032	0,037

Tabela C.6: Absolutna pristranskost zaradi nepokritja za spremenljivko spol

Lestvica: 0 (Moški), 1 (Ženski)

Negativni predznak: vzorčni okvir izključno uporabnikov interneta precenjuje delež moških.

Država/Leto	2010	2011	2012	2013	2014	2015	2016	Povprečje
Avstrija	-4,2 %	-4,1 %	-2,2 %	-1,9 %	-1,7 %	-1,3 %	-1,1 %	-2,5 %
Belgija	-2,1 %	-1,2 %	-0,9 %	-0,9 %	-2,8 %	-0,8 %	-1,0 %	-1,4 %
Bolgarija	0,1 %	-2,9 %	-0,8 %	0,9 %	-2,8 %	-1,3 %	-1,8 %	-1,2 %
Hrvaška	0,3 %	-0,3 %	-0,4 %	-0,2 %	-2,3 %	-1,5 %	-1,1 %	-0,8 %
Republika Ciper	-0,7 %	-1,1 %	-0,1 %	-1,7 %	-1,6 %	-2,1 %	-1,8 %	-1,3 %
TCC		-0,4 %	-2,9 %	-0,2 %	-0,7 %	-3,3 %	-0,8 %	-1,4 %
Češka	-3,4 %	-3,4 %	-2,5 %	-3,3 %	-3,8 %	-1,9 %	-0,5 %	-2,7 %
Danska	-1,3 %	-0,6 %	-0,9 %	-0,5 %	-0,7 %	-0,4 %	-0,6 %	-0,7 %
Estonija	-4,5 %	-2,0 %	-2,4 %	-1,2 %	-1,3 %	-1,8 %	-3,7 %	-2,4 %
Finska	-2,0 %	-1,6 %	-2,2 %	-1,8 %	-1,4 %	-2,3 %	-0,8 %	-1,7 %
Francija	-3,2 %	-1,9 %	-2,4 %	-1,2 %	-1,3 %	0,2 %	-0,4 %	-1,5 %
Nemčija – Vzhod	-4,2 %	-4,5 %	-1,9 %	-1,2 %	-0,7 %	-4,0 %	-1,8 %	-2,6 %
Nemčija – Zahod	-2,5 %	-2,6 %	-2,0 %	-2,2 %	-1,7 %	-2,1 %	-2,2 %	-2,2 %
Velika Britanija	-2,1 %	-2,4 %	-0,5 %	-1,6 %	-1,8 %	-1,1 %	-0,7 %	-1,5 %
Grčija	-4,7 %	-2,2 %	-3,6 %	-5,0 %	-2,5 %	-3,1 %	-0,7 %	-3,1 %
Madžarska	-2,0 %	-2,6 %	-2,3 %	0,8 %	-0,9 %	-3,0 %	-2,7 %	-1,8 %
Irska	-0,1 %	-0,4 %	0,2 %	1,6 %	-0,5 %	0,4 %	-0,1 %	0,2 %
Italija	-2,9 %	0,1 %	1,3 %	-1,5 %	-5,4 %	-3,5 %	-1,0 %	-1,8 %
Latvija	-0,3 %	-1,7 %	-0,6 %	0,3 %	-0,8 %	-1,7 %	-2,0 %	-1,0 %
Litva	-1,7 %	0,5 %	-3,0 %	-2,5 %	-2,2 %	-2,4 %	-2,5 %	-2,0 %
Luksemburg	-3,2 %	-2,7 %	-2,0 %	-0,5 %	-0,7 %	-3,1 %	-0,4 %	-1,8 %
Malta	0,8 %	-1,9 %	-1,6 %	-2,6 %	-2,8 %	-2,3 %	-0,8 %	-1,6 %
Nizozemska	-0,7 %	0,0 %	-0,2 %	-0,3 %	-0,5 %	-0,3 %	0,1 %	-0,3 %
Severna Irska	-2,9 %	1,0 %	2,7 %	-4,9 %	0,3 %	-2,2 %	-0,4 %	-0,9 %
Poljska	-1,5 %	-1,6 %	-0,8 %	-1,5 %	-2,0 %	-1,5 %	-1,6 %	-1,5 %
Portugalska	-2,8 %	-2,1 %	0,1 %	-1,1 %	-2,2 %	-1,4 %	-1,3 %	-1,5 %
Romunija	-2,1 %	-0,8 %	-0,4 %	-0,6 %	-2,6 %	-3,4 %	-1,4 %	-1,6 %
Slovaška	-4,7 %	-2,2 %	-0,9 %	-2,7 %	-1,4 %	-2,5 %	-2,1 %	-2,4 %
Slovenija	-2,2 %	-2,2 %	-1,1 %	-2,1 %	-2,0 %	-2,5 %	-1,9 %	-2,0 %
Španija	-2,1 %	-1,8 %	-2,0 %	-1,3 %	-4,4 %	-2,5 %	-1,4 %	-2,2 %
Švedska	-1,2 %	-1,2 %	-0,4 %	-0,5 %	-1,7 %	-1,0 %	-2,1 %	-1,2 %
Turčija	-3,6 %	-2,8 %	-3,3 %	2,9 %	-3,7 %	2,3 %	-1,9 %	-1,5 %
EU27	-2,4 %	-1,8 %	-1,2 %	-1,5 %	-2,3 %	-1,7 %	-1,2 %	-1,7 %
EU28+TUR+TCC		-1,8 %	-1,2 %	-1,0 %	-2,4 %	-1,3 %	-1,1 %	-1,5 %

Tabela C.7: Absolutna pristranskost zaradi nepokritja za spremenljivko starost

Negativni predznak: vzorčni okvir izključno uporabnikov interneta podcenjuje povprečno starost

Država/Leto	2010	2011	2012	2013	2014	2015	2016	Povprečje
Avstrija	-6.071	-6.365	-6.044	-5.450	-5.650	-6.081	-4.650	-5.759
Belgija	-6.071	-6.365	-6.044	-5.450	-5.650	-6.081	-4.650	-5.759
Bolgarija	-9.164	-8.923	-7.770	-7.681	-9.376	-8.753	-8.746	-8.630
Hrvaška	-7.138	-7.011	-7.158	-7.058	-6.397	-5.256	-4.650	-6.381
Republika Ciper	-7.757	-7.590	-6.051	-4.680	-4.956	-6.226	-5.365	-6.089
TCC		-5.095	-3.907	-4.358	-3.250	-4.399	-1.460	-3.745
Češka	-6.903	-6.780	-5.719	-5.120	-5.400	-5.639	-4.660	-5.746
Danska	-2.989	-2.641	-1.696	-1.684	-1.473	-1.229	-2.141	-1.979
Estonija	-6.194	-7.376	-5.741	-5.480	-4.888	-5.275	-5.845	-5.828
Finska	-4.134	-3.512	-4.112	-3.099	-3.524	-2.672	-2.938	-3.427
Francija	-5.304	-4.495	-3.629	-3.882	-3.937	-4.136	-3.156	-4.077
Nemčija - Vzhod	-6.129	-6.268	-5.944	-6.762	-5.246	-6.235	-4.211	-5.828
Nemčija - Zahod	-6.054	-5.284	-4.022	-4.040	-4.561	-3.747	-3.547	-4.465
Velika Britanija	-4.185	-3.598	-3.979	-3.041	-2.012	-3.099	-2.817	-3.247
Grčija	-9.204	-8.337	-7.279	-8.492	-6.813	-7.754	-6.732	-7.802
Madžarska	-7.039	-6.377	-6.170	-5.599	-5.828	-6.528	-5.853	-6.199
Irska	-3.477	-2.994	-3.206	-3.187	-2.842	-2.753	-2.651	-3.016
Italija	-4.326	-3.984	-4.991	-5.194	-5.042	-6.669	-6.401	-5.230
Latvija	-4.860	-4.996	-4.241	-3.761	-4.908	-5.041	-4.195	-4.572
Litva	-9.186	-8.818	-7.921	-7.492	-7.272	-7.217	-7.064	-7.853
Luksemburg	-3.298	-2.990	-2.896	-2.853	-1.344	-2.367	-1.828	-2.511
Malta	-4.782	-5.478	-4.804	-4.250	-4.404	-4.014	-3.855	-4.512
Nizozemska	-1.410	-0.856	-1.186	-0.931	-0.759	-0.562	-0.342	-0.864
Severna Irska	-4.474	-4.076	-4.365	-3.899	-2.897	-4.125	-2.295	-3.733
Poljska	-6.837	-7.168	-6.446	-6.620	-5.660	-6.179	-6.388	-6.471
Portugalska	-7.398	-9.399	-9.451	-9.774	-8.403	-7.006	-7.485	-8.416
Romunija	-8.543	-7.611	-7.912	-7.188	-7.048	-8.065	-8.193	-7.794
Slovaška	-6.981	-5.771	-5.277	-5.124	-5.318	-5.198	-5.581	-5.607
Slovenija	-6.192	-6.340	-5.875	-5.056	-4.710	-5.484	-3.900	-5.365
Španija	-8.006	-6.713	-6.807	-5.959	-5.348	-6.471	-5.577	-6.411
Švedska	-2.309	-2.042	-1.597	-1.088	-1.796	-0.889	-1.887	-1.658
Turčija	-4.290	-4.524	-5.003	-5.127	-4.133	-4.838	-2.079	-4.285
EU27	-5.375	-4.910	-4.661	-4.546	-4.356	-4.739	-4.358	-4.707
EU28+TUR+TCC		-4.557	-4.304	-4.344	-4.123	-4.459	-4.112	-4.317

Tabela C.8: Absolutna pristranskost zaradi nepokritja za spremenljivko starost pri končani izobrazbi

Pozitivni predznak: vzorčni okvir izključno uporabnikov interneta precenjuje povprečno starost pri končani izobrazbi

Država/Leto	2010	2011	2012	2013	2014	2015	2016	Povprečje
Avstrija	0,582	0,655	0,601	0,453	0,769	0,511	0,586	0,594
Belgija	0,582	0,655	0,601	0,453	0,769	0,511	0,586	0,594
Bolgarija	1,484	1,089	1,395	1,135	1,397	1,029	1,159	1,241
Hrvaška	1,166	0,894	0,952	0,712	0,597	0,414	0,586	0,760
Republika Ciper	1,985	1,749	1,697	0,942	1,511	1,402	1,185	1,496
TCC		0,941	1,037	1,032	0,657	0,998	0,215	0,813
Češka	0,512	0,461	0,482	0,439	0,432	0,481	0,387	0,456
Danska	0,526	0,398	0,302	0,317	0,150	0,008	0,144	0,264
Estonija	0,651	0,469	0,410	0,431	0,522	0,505	0,590	0,511
Finska	0,813	0,724	0,652	0,511	0,747	0,537	0,934	0,703
Francija	0,955	0,821	0,750	0,720	0,679	0,627	0,378	0,704
Nemčija - Vzhod	0,966	0,739	0,638	0,739	0,701	0,914	0,746	0,778
Nemčija - Zahod	0,834	0,586	0,521	0,681	0,622	0,512	0,664	0,632
Velika Britanija	0,552	0,546	0,545	0,494	0,458	0,386	0,352	0,476
Grčija	2,654	2,284	2,008	1,737	1,855	1,883	1,653	2,010
Madžarska	1,324	1,244	1,137	0,986	1,070	0,773	0,837	1,053
Irska	0,660	0,505	0,618	0,705	0,493	0,379	0,468	0,547
Italija	1,468	1,409	1,439	1,518	1,559	1,914	1,612	1,560
Latvija	0,486	0,618	0,425	0,343	0,543	0,412	0,438	0,467
Litva	1,081	1,089	0,775	0,978	0,839	0,714	0,948	0,918
Luksemburg	0,555	0,529	0,420	0,403	0,448	0,528	0,432	0,474
Malta	0,559	0,643	0,581	0,596	0,580	0,619	0,449	0,576
Nizozemska	0,201	0,195	0,146	0,144	0,159	0,148	0,085	0,154
Severna Irska	0,649	0,674	0,610	0,588	0,342	0,478	0,461	0,543
Poljska	0,987	1,150	1,122	0,911	0,918	0,801	0,816	0,958
Portugalska	2,455	1,933	2,148	2,053	1,570	1,666	1,637	1,923
Romunija	2,070	1,879	1,743	1,274	1,621	1,058	1,211	1,551
Slovaška	0,758	0,560	0,535	0,475	0,606	0,521	0,445	0,557
Slovenija	0,925	0,876	0,740	0,618	0,528	0,549	0,611	0,692
Španija	1,736	1,675	1,466	1,300	1,111	1,258	1,286	1,405
Švedska	0,545	0,314	0,267	0,166	0,399	-0,005	0,035	0,246
Turčija	1,215	1,635	1,362	1,315	1,329	1,327	1,280	1,352
EU27	1,179	1,058	1,013	0,967	0,923	0,880	0,872	0,985
EU28+TUR+TCC		1,233	1,210	1,101	1,027	1,010	0,864	1,074

Tabela C.9: Absolutna pristranskost zaradi nepokritja za spremenljivko *politična usmeritev levo-desno*

Lestvica: 1 (levo) – 10 (desno)

Negativni predznak: vzorčni okvir izključno uporabnikov interneta precenjuje levo politično usmerjenost

Država/Leto	2010	2011	2012	2013	2014	2015	2016	Povprečje
Avstrija	-0,028	0,080	-0,011	-0,070	-0,081	-0,080	-0,149	-0,048
Belgija	-0,028	0,080	-0,011	-0,070	-0,081	-0,080	-0,149	-0,048
Bolgarija	0,552	0,524	0,559	0,551	0,807	0,874	1,012	0,697
Hrvaška	0,001	-0,069		-0,058	-0,032	-0,242	-0,149	-0,092
Republika Ciper	0,168	0,163	-0,201	0,152	0,034	0,036	-0,048	0,043
TCC		-0,398			-0,121	-0,209	-0,165	-0,223
Češka	0,419	0,529	0,397	0,315	0,347	0,366	0,312	0,384
Danska	-0,097	0,006	-0,026	-0,021	-0,001	-0,032	-0,036	-0,030
Estonija	0,105	0,100	0,037	0,098	0,027	0,110	0,061	0,077
Finska	0,011	0,025	0,027	0,032	0,033	0,030	-0,014	0,021
Francija	-0,082	-0,046	0,004	-0,079	-0,069	-0,019	-0,093	-0,055
Nemčija - Vzhod	-0,009	-0,100	0,083	-0,085	0,086	0,042	-0,045	-0,004
Nemčija - Zahod	-0,142	-0,130	-0,056	-0,103	-0,090	-0,064	-0,088	-0,096
Velika Britanija	0,014	0,005	-0,046	-0,063	-0,004	-0,049	0,009	-0,019
Grčija	-0,289	-0,226	-0,211	-0,148	-0,103	-0,227	-0,041	-0,178
Madžarska	0,214	0,225	0,241	0,092	0,165	0,350	0,100	0,198
Irska	-0,060	-0,124	-0,067	0,110	-0,072	-0,058	0,033	-0,034
Italija	0,059	0,140	0,079	0,075	0,096	-0,307	-0,075	0,009
Latvija	0,102	0,027	0,022	-0,050	-0,005	0,050	0,055	0,029
Litva	-0,136	0,061	-0,053	0,110	-0,049	-0,009	0,187	0,016
Luksemburg	-0,127	-0,096	-0,040	-0,014	-0,059	-0,063	-0,008	-0,058
Malta	-0,115	0,117	0,082	-0,046	0,118	0,025	0,052	0,033
Nizozemska	-0,024	-0,020	-0,001	-0,054	0,004	-0,015	-0,006	-0,017
Severna Irska	-0,034	-0,137	-0,074	-0,035	-0,119	0,014	-0,019	-0,058
Poljska	-0,046	-0,010	-0,040	0,013	-0,080	-0,228	-0,084	-0,068
Portugalska	-0,106	0,082	-0,018	0,138	0,126	0,071	0,118	0,059
Romunija	0,202	0,213	0,225	0,121	0,267	0,425	0,236	0,241
Slovaška	0,180	0,250	0,231	0,201	0,236	0,175	0,395	0,238
Slovenija	0,022	0,017	-0,043	-0,043	-0,071	-0,062	-0,115	-0,042
Španija	-0,084	-0,013	-0,233	-0,031	-0,101	-0,211	-0,105	-0,111
Švedska	0,016	-0,004	-0,002	-0,023	0,044	0,037	0,022	0,013
Turčija	-0,419	-0,295			-0,100	-0,226	-0,155	-0,239
EU27	-0,039	-0,002	-0,006	-0,015	-0,017	-0,059	-0,027	-0,024
EU28+TUR+TCC		-0,090			-0,038	-0,109	-0,037	-0,068

Tabela C.10: Absolutna pristranskost zaradi nepokritja za spremenljivko zadovoljstvo z življenjem

Lestvica: 1 (sploh ni zadovoljen/zadovoljna) – 4 (zelo zadovoljen/zadovoljna)

Pozitivni predznak: vzorčni okvir izključno uporabnikov interneta precenjuje povprečno zadovoljstvo

Država/Leto	2010	2011	2012	2013	2014	2015	2016	Povprečje
Avstrija	0,076	0,064	0,037	0,052	-0,028	0,038	0,055	0,042
Belgija	0,076	0,064	0,037	0,052	-0,028	0,038	0,055	0,042
Bolgarija	0,298	0,232	0,207	0,195	0,200	0,169	0,168	0,210
Hrvaška		0,191	0,159	0,141	0,107	0,082	0,055	0,122
Republika Ciper	0,067	0,156	0,122	0,052	0,101	0,080	0,075	0,093
TCC		0,071	0,052	0,068	0,031	0,026	0,027	0,046
Češka	0,123	0,109	0,069	0,086	0,072	0,076	0,058	0,085
Danska	0,024	0,027	0,014	0,021	0,003	0,002	0,015	0,015
Estonija	0,057	0,088	0,090	0,070	0,080	0,068	0,084	0,077
Finska	0,035	0,016	0,027	0,030	0,021	0,011	0,025	0,023
Francija	0,075	0,039	0,062	0,059	0,064	0,044	0,045	0,055
Nemčija - Vzhod	0,050	0,073	0,044	0,082	0,060	0,045	0,045	0,057
Nemčija - Zahod	0,070	0,021	0,053	0,030	0,028	0,021	0,047	0,039
Velika Britanija	0,029	0,056	0,028	0,024	0,027	0,026	0,029	0,031
Grčija	0,174	0,238	0,138	0,114	0,138	0,103	0,092	0,143
Madžarska	0,154	0,172	0,110	0,114	0,151	0,092	0,109	0,129
Irska	0,038	0,021	0,026	0,055	0,022	0,033	0,024	0,031
Italija	0,110	0,079	0,073	0,045	0,070	0,100	0,042	0,074
Latvija	0,131	0,154	0,093	0,086	0,092	0,080	0,085	0,103
Litva	0,185	0,171	0,125	0,114	0,129	0,110	0,098	0,133
Luksemburg	0,012	0,016	-0,014	0,019	-0,013	0,035	0,002	0,008
Malta	0,020	0,091	0,042	0,071	0,012	0,019	0,012	0,038
Nizozemska	0,024	0,018	0,009	0,001	0,006	0,007	0,003	0,010
Severna Irska	0,018	0,049	0,044	-0,026	0,032	-0,007	0,028	0,020
Poljska	0,123	0,107	0,100	0,111	0,085	0,079	0,110	0,102
Portugalska	0,189	0,162	0,155	0,153	0,165	0,063	0,118	0,144
Romunija	0,255	0,223	0,208	0,204	0,138	0,161	0,182	0,196
Slovaška	0,125	0,099	0,150	0,095	0,103	0,071	0,069	0,102
Slovenija	0,069	0,101	0,048	0,051	0,055	0,053	0,075	0,064
Španija	0,085	0,085	0,111	0,110	0,106	0,069	0,093	0,094
Švedska	0,030	0,010	0,002	0,011	0,005	0,015	0,005	0,011
Turčija		0,052	0,059	-0,008	0,042	-0,024	-0,022	0,017
EU27	0,139	0,123	0,124	0,109	0,102	0,101	0,097	0,114
EU28+TUR+TCC		0,131	0,126	0,103	0,103	0,097	0,095	0,109