

UNIVERZA V LJUBLJANI
FAKULTETA ZA DRUŽBENE VEDE

JANA ŠTRAKL

VEČNIVOJSKA ANALIZA

DIPLOMSKO DELO

LJUBLJANA 2008

UNIVERZA V LJUBLJANI
FAKULTETA ZA DRUŽBENE VEDE

JANA ŠTRAKL
MENTOR: DOC. DR. GREGOR PETRIČ

VEČNIVOJSKA ANALIZA

DIPLOMSKO DELO

LJUBLJANA 2008

ZAHVALA

*Hvala vsem, ki ste skupaj z mano hodili po poti do cilja.
Hvala mentorju doc. dr. Gregorju Petriču za strokovno podporo, prijateljstvo in
za pomoč pri realizaciji tega dela.
Zahvaljujem se svoji družini ter prijateljem za moralno podporo, brez katerih
študentska leta ne bi bila tako posebna.
Blažu za tolažbo in pozitivno energijo, ki jo je nesebično delil z mano in mi nikoli
ni pustil obupati.*

VEČNIVOJSKA ANALIZA

Na veliko različnih raziskovalnih področij se srečamo s hierarhično podatkovnimi strukturami, a do nedavnega ni obstajala primerna tehnika za analiziranje takšnih vrst podatkov. Gre za podatke z združenim virom variabilnosti, torej podatke, kjer se posamezniki združujejo glede na skupine, katerim pripadajo. Takšne vrste podatkov najdemo v različnih raziskovalnih vedah, saj ljudje (ter tudi druga živa bitja) obstajamo znotraj različnih organizacijskih struktur, kot so družine, šole, mesta in podobno. Analiza, ki omogoča analiziranje hierarhičnih podatkov, se imenuje večnivojska analiza, ki je veliko naprednejša oblika enostavne linearne regresije in multiple linearne regresije. Večnivojska analiza namreč dovoljuje, da varianco v odvisnih spremenljivkah analiziramo na multiplih hierarhičnih nivojih, medtem ko so v regresiji vsi učinki modelirani le na enem nivoju. V diplomskem delu sem se ukvarjala s prednostmi večnivojske analize v teoretičnem in empiričnem delu in jo primerjala z regresijsko analizo, ki je bila v preteklosti edina metoda, s katero smo raziskovali takšne vrste podatkov.

Ključne besede: večnivojska analiza, regresija, deagregacija, hierarhične podatkovne strukture.

MULTILEVEL ANALYSIS

In many different areas of research we can recognize the existence of hierarchical data structure and until recently there has not been an appropriate technique for analysis of these types of data. These data have nested source of variability, implying that we have individuals nested into higher level unit presented by groups. This kind of data can be found in different areas as people (and other living beings) tend to exist within different organizational structures like families, schools and towns among many others. Analysis that allows examination of hierarchical data is multilevel analysis, which is an advanced form of ordinary linear regression and multiple regression analysis. Multilevel analysis allows to examine variance in outcome at multiple hierarchical levels, meanwhile in regression all effects are modelled at one level. In this paper I am dealing with advantages of multilevel modelling, theoretically and empirically, and comparing it with regression analysis, which was in the past the only method employed to examine this kinds of data.

Keywords: multilevel analysis, regression, disaggregation, hierarchical data structure.

KAZALO

1. UVOD	9
1.1 CILJI NALOGE.....	10
2. VEČNIVOJSKA ANALIZA	13
3. PODATKOVNE STRUKTURE	17
4. PREDNOSTI VEČNIVOJSKE ANALIZE	21
4.1 TRADICIONALNI PRISTOPI REGRESIJSKEGA MODELIRANJA VEČNIVOJSKIH PODATKOV.....	22
4.2 PRAKTIČNI PRIMER – UPORABNOST VEČNIVOJSKEGA MODELA.....	25
5. PREDPOSTAVKE REGRESIJE IN VEČNIVOJSKE ANALIZE	28
5.1 VZORČENJE.....	28
5.2 NEODVISNOST OPAZOVANJ.....	29
5.3 LINEARNOST.....	30
5.4 HOMOSCEDASTIČNOST.....	30
5.5 NORMALNOST PORAZDELITVE REZIDUALOV.....	31
6. STATISTIČNO OZADJE VEČNIVOJSKE ANALIZE	32
6.1 RAZISKAVA POVEZAVE MED DRUŽBENO EKONOMSKIM STATUSOM IN USPEHOM V ŠOLI.....	32
6.2 RAZISKAVA POVEZAVE MED DRUŽBENO EKONOMSKIM STATUSOM V DVEH ŠOLAH.....	34
6.3 RAZISKAVA POVEZAVE MED DRUŽBENO EKONOMSKIM STATUSOM IN USPEHOM V J ŠOLAH.....	35
7. RAZISKOVALNI OKVIR IN ANALIZA	42
7.1 POJASNJEVALNI MODEL.....	42
7.2 OPIS VZORCA IN ZBIRANJE PODATKOV.....	43
7.3 INDIKATORJI.....	47
7.4 PRIMERJAVA OPISNIH STATISTIK.....	49
7.5 REZULTATI REGRESIJE S FIKSNIMI UČINKI.....	51
7.6 REZULTATI VEČNIVOJSKE ANALIZE.....	53
7.7 PRIMERJAVA DVEH ANALITIČNIH PRISTOPOV.....	57
8. ZAKLJUČEK	61

9. LITERATURA.....	64
10. PRILOGE.....	67
PRILOGA A: REZULTATI VEČNIVOJSKE ANALIZE.....	67
PRILOGA B: REZULTATI REGRESIJE S FIKSNIMI UČINKI.....	71

KAZALO TABEL

Tabela 7.1: OPIS VZORCA IN STRUKTURA UPORABNIKOV – SOCIODEMOGRAFSKI PODATKI.....	43
Tabela 7.2: OPISI ODVISNE IN NEODVISNIH SPREMENLJIVK UPORABLJENIH V RAZISKAVI.....	47
Tabela 7.3: OPISNE STATISTIKE V REGRESIJI.....	49
Tabela 7.4: OPISNE STATISTIKE HIERARHIČNIH PODATKOV ZA NIVO 2.....	49
Tabela 7.5: OPISNE STATISTIKE HIERARHIČNIH PODATKOV ZA NIVO 1.....	50
Tabela 7.6: REGRESIJSKI KOEFICIENTI.....	51
Tabela 7.7: POVZETEK MODELA.....	52
Tabela 7.8: OCENE DELEŽEV NAKLJUČNIH UČINKOV.....	54
Tabela 7.9: KOEFICIENTI.....	55
Tabela 7.10: PRIMERJAVA DVEH STATISTIČNIH PRISTOPOV.....	58

KAZALO SLIK

Slika 3.1: DIAGRAM ENOT DVONIVOJSKE HIERARHIJE.....	18
Slika 3.2: KLASIFIKACIJSKI DIAGRAM DVONIVOJSKE HIERARHIJE.....	18
Slika 3.3: DIAGRAM TRINIVOJSKE HIERARHIJE.....	18
Slika 3.4: KLASIFIKACIJSKI DIAGRAM TRINIVOJSKE HIERARHIJE.....	19
Slika 3.5: DIAGRAM NAVZKRIŽNE KLASIFIKACIJE.....	20
Slika 3.6: KLASIFIKACIJSKI DIAGRAM NAVZKRIŽNE KLASIFIKACIJE.....	20
Slika 4.1: GRAFIČNI PRIKAZ ENOSTAVNEGA PRIMERA.....	25
Slika 5.1: RAZLIKA MED ENOSTAVNIM SLUČAJNIM VZORČENJEM IN DVOSTOPENJSKIM VZORČENJEM.....	29
Slika 5.2: PRIMER KRIVULJASTE LINEARNE IN LINEARNE POVEZAVE.....	30
Slika 5.3: GRAFIČNI PRIKAZ HOMOSCEDASTIČNOSTI IN HETEROSCEDASTIČNOSTI.....	31
Slika 6.1: POVEZAVE MED USPEHOM IN DRUŽBENO EKONOMSKIM STATUSOM V ENI HIPOTETIČNI ŠOLI.....	32
Slika 6.2: POVEZAVA MED USPEHOM IN CENTRIRANO SPREMENLJIVKO DRUŽBENO EKONOMSKI STATUS V HIPOTETIČNI ŠOLI.....	33
Slika 6.3: POVEZAVA MED USPEHOM IN DRUŽBENO EKONOMSKIM STATUSOM ZNOTRAJ DVEH HIPOTETIČNIH ŠOL.....	34
Slika 6.4: PRIKAZ POVPREČJA ŠOL (VERTIKALNO) IN NAGIBOV DRUŽBENO EKONOMSKEGA STATUSA (HORIZONTALNO) ZA 200 HIPOTETIČNIH ŠOL.....	37
Slika 6.5: GRAF POVPREČIJ ŠOL IN NAGIBOV DRUŽBENO EKONOMSKEGA STATUSA ZA 100 HIPOTETIČNIH KATOLIŠKIH ŠOL.....	39
Slika 6.6: GRAF POVPREČIJ ŠOL IN NAGIBOV DRUŽBENO EKONOMSKEGA STATUSA ZA 100 HIPOTETIČNIH JAVNIH ŠOL.....	40
Slika 7.1: FREKVENCE FORUMOV.....	45
Slika 7.2: DIAGRAM ENOT.....	46
Slika 7.3: KLASIFIKACIJSKI DIAGRAM.....	46

1. UVOD

Ideja, da družbeni konteksti vplivajo na posameznikovo delovanje, je osnovna ideja sociologije kot znanosti. Za sociologa je torej ključno vprašanje, kako skupine ali organizacije vplivajo na delovanje posameznika, kako se posamezniki znotraj teh kontekstov obnašajo in kako se poistovetijo s pripadniki istega družbenega konteksta. Gledano z vidika statističnih raziskovanj takšnih vprašanj je imel raziskovalec še ne dolgo nazaj na razpolago precej omejene vrste statističnih analiz. Na vprašanja interakcij in medsebojnih vplivov posameznikov in širših družbenih kontekstov, v katerih se le-ti nahajajo, ni bilo mogoče odgovoriti. Znanstveniki so torej imeli nemalo problemov pri poudarjanju pomembnih vplivov skupin na posameznike in le-tega niso mogli potrditi s statistično analizo. V zadnjem desetletju so v ta namen strokovnjaki razvili za sociologijo izjemno pomembno statistično analizo, imenovano večnivojska analiza. Zagovarjajo namreč pomembnost upoštevanja vplivov skupin pri raziskovanju posameznikovih delovanj, saj so enotnega mnenja, da neupoštevanje le-teh vodi v zmotne interpretacije rezultatov.

S spreminjanjem družbe in z vse večjo nujo po razumevanju različnih družbenih fenomenov se v raziskovanju posameznikov, družbenih struktur in trgov vzporedno torej spreminjajo tudi načini raziskovanja. Statistične analize so k raziskovanju in razumevanju sveta pripomogle veliko, saj z njimi lahko razkrijemo različne povezave med posamezniki in družbenimi strukturami, vendar s tem potrebe še zdaleč niso zadovoljene. Ravno nasprotno, strokovnjaki vsak dan razvijajo nove metode in analize na različnih področjih raziskovanja. Uporabljajo se pravzaprav v vseh znanstvenih vedah, saj za strokovnjake predstavljajo eno najpomembnejših raziskovalnih orodij. Zasedimo jih tako v sociologiji, ki se ukvarja z osnovnim vprašanjem povezave med posamezniki in družbenimi strukturami, v ekonomiji, kjer je zanimanje usmerjeno na različne trge, v psihologiji oziroma raziskovanju posameznika, v biologiji, kjer je raziskovanje usmerjeno v raziskovanje živalskega in rastlinskega sveta, in še bi lahko naštevali. Različne statistične metode so na

vseh teh področjih nepogrešljivo orodje pri iskanju odgovorov na raziskovalna vprašanja.

V poplavi različnih statističnih analiz je tako osnovno vprašanje raziskovalca, katera metoda se najbolje prilega njegovemu raziskovalnemu ukroju. V diplomskem delu se bom ukvarjala z dvema tipoma na videz podobnih statističnih analiz in poskušala ugotoviti, katera analiza je glede na vrsto podatkov in tip raziskovalnega vprašanja boljša, in ugotavljala, s katero lahko bolje pojasnimo družbene strukture in posameznike.

1.1 CILJI NALOGE

V nalogi želim predstaviti eno izmed najbolj perspektivnih statističnih metod – večnivojsko analizo –, ki je pri nas v veliki meri neizkoriščena. Velikokrat namreč preučujemo enote, ki so lahko v širšem kontekstu del neke višje enote, saj se enote združujejo v različne, a kontekstualno enake skupine. Te skupine lahko predstavljajo na različnih področjih različne vrste skupin, kot denimo šole na področju izobraževanja ali mesta, pokrajine in države na področju geografskega preučevanja. V sociologiji se srečujemo z izjemnim številom različnih skupin, v katere so posamezniki vključeni, v njih sodelujejo, so njihovi člani. Ravno preučevanje posameznika v tem širšem kontekstu je eden izmed pglavitnih elementov sociologije kot vede. Skupine so namreč sestavljene iz posameznikov, ki slonijo na določeni vrsti vzajemnega delovanja. Skupine lahko močno vplivajo na posameznike, ki so njihovi člani in obratno; posamezniki, ki to skupino sestavljajo, vplivajo na lastnosti te skupine, jo oblikujejo, spreminjajo itd. Seveda pa se je potrebno zavedati, da so skupine znotraj enakega družbenega okvirja med seboj lahko zelo različne in zato notranje bolj homogene. Če za primer vzamemo področje šolstva in šole kot skupine, lahko ugotovimo, da različne šole privlačijo različne učence. Nekatere lahko privlačijo učence iz bogatih družin, druge iz revnih, nekatere šole lahko privlačijo bolj verne učence, druge manj verne itd. Gre torej za preučevanje interakcij med skupinami in njenimi pripadniki ter vplive obeh elementov na zastavljeno raziskovalno vprašanje. V večnivojski analizi sta tako pomembna oba faktorja oziroma nivoja pri preučevanju navzkrižno nivojskih hipotez oziroma

večnivojskega problema. Večnivojski problem je torej problem, ki se nanaša na povezavo spremenljivk, merjenih na različnih hierarhičnih nivojih. Najpogostejše vprašanje, ki se v tem primeru zastavlja, je, kako spremenljivke višjega nivoja in spremenljivke nižjega nivoja vplivajo na raziskovano spremenljivko nižjega nivoja. Cilj analize je torej določiti neposredne vplive pojasnjevalnih spremenljivk nižjega in višjega nivoja in določiti ali spremenljivke skupinskega nivoja služijo kot razsodniki povezav na nižjem nivoju. Če se izkaže, da spremenljivke višjih nivojev moderirajo povezave na nižjem nivoju, se pojavlja statistična interakcija med pojasnjevalnimi spremenljivkami na različnih nivojih. Za pojasnjevanje vplivov se v največji meri uporablja regresijska analiza, vendar je njena temeljna slabost, da lahko večnivojske podatke obravnava samo na istem nivoju (preko agregacije ali deagregacije), kar pa vnaša statistične težave, kot bomo videli kasneje.

V diplomski nalogi bom na teoretični in empirični ravni skušala predstaviti prednosti večnivojske analize pred regresijsko analizo. Za predstavitev primerjave se bom velikokrat sklicevala na primer šol ter učencev, ki je v tovrstni literaturi najbolj priljubljen primer. V drugem, praktičnem delu naloge bom uporabila obe metodi na primeru pojasnjevalnega modela reflektivnost v spletnih forumih (Petrič 2007) in podatkih, ki so bili zbrani v namen preverjanja tega modela v okviru raziskovalnega projekta "Družbeni potencial slovenskih spletnih forumov", ki se je izvajal na Centru za metodologijo in informatiko na FDV. Relevantnost teh podatkov za naš problem izhaja iz dejstva, da združuje posameznike tako na nižjem nivoju, v našem primeru so to uporabniki forumov, ki torej v večnivojski analizi predstavljajo nižji nivo, in na višjem nivoju, ki v našem primeru predstavlja različne forume. Z obema analizama bom skušala pojasniti delovanje posameznikov na različnih forumih. Pri tem bom uporabila statistična paketa SPSS za regresijsko analizo ter HLM za večnivojsko analizo.

Diplomska naloga je razdeljena na devet poglavij. Prvih šest poglavij je namenjenih teoretični razpravi in pojasnjevanju pomembnih elementov večnivojske analize, sledi empirični del naloge ter nazadnje še poglavje z ugotovitvami. Uvodu torej sledi poglavje o razumevanju večnivojske analize ter njen razvoj. V tretjem poglavju se bom ukvarjala s podatkovnimi strukturami,

kjer bom poskušala prikazati primerne vrste podatkov, ki jih s to vrsto analize lahko analiziramo ter prikazati pomembnost različnih struktur. V četrtem poglavju bom poskušala predstaviti prednosti večnivojskega modeliranja v primerjavi z metodami, ki se običajno uporabljajo v pojasnjevalne namene. V petem poglavju bom predstavila in poudarila pomembnost predpostavk in kršenja le-teh. Sledi šesto poglavje, kjer bom poskušala nakazati statistično ozadje večnivojske analize na primeru, ki se v tovrstni literaturi največkrat pojavlja. Sedmo poglavje je namenjeno praktičnemu delu naloge, kjer bom na podatkih o spletnih forumih in njihovih uporabnikih (Petrič 2007) izvedla večnivojsko analizo ter multiplo regresijsko analizo ter ju med seboj primerjala. Osmo poglavje je namenjeno ugotovitvam in sklepom, pridobljenih iz celotnega dela.

2. VEČNIVOJSKA ANALIZA

Pojem »večnivojsko« se nanaša na podatke, ki imajo združeni vir variabilnosti, torej podatke, ki vsebujejo enote na nižjem nivoju, imenovanem tudi mikro nivo, in so povezani v višji nivo, imenovanem tudi makro nivo. Večnivojska analiza nam omogoča sočasno raziskavo učinkov višjega nivoja (t.i. skupinskega nivoja saj se enote združujejo glede na skupine) in nižjega nivoja (t.i. nivo posameznikov) z upoštevanjem neodvisnosti merjenja znotraj skupin. Omogoča nam tudi raziskavo variabilnosti med skupinami ter znotraj skupin in pojasnitev, kako so spremenljivke višjega in nižjega nivoja povezane z variabilnostjo na obeh nivojih. Na ta način lahko uporabimo večnivojske modele za sklepanje o povzročiteljih variacije med posamezniki ali o povezavi neodvisnih spremenljivk višjega in nižjega nivoja z obnašanjem posameznikov. Sklepamo lahko tudi o variaciji med skupinami, če le-ta v podatkih obstaja, in v kakšni meri jo pojasnjujejo lastnosti obeh nivojev. Pomembno je tudi dejstvo, da skupine niso privzete kot nepovezane, ampak so obravnavane kot izhajajoče iz večje populacije skupin, katero želimo raziskovati. Če povzamemo, pri večnivojski analizi se torej raziskovalec lahko ukvarja tako z mikro nivojem posameznikov in makro nivojem skupin kot tudi z obojim hkrati (Hox 2002).

Pomembno je dejstvo, da večnivojski modeli temeljijo na večnivojski teoriji, ki podrobno razlaga pričakovane neposredne učinke spremenljivk ene na drugo znotraj katerega koli nivoja ter učinke navzkrižno-nivojskih interakcij med spremenljivkami, ki se nahajajo na različnih nivojih. Raziskovalec mora torej zagotavljati neposredne mehanizme, ki povzročajo vplive spremenljivk na enem nivoju na spremenljivke na drugem nivoju. Večnivojska analiza torej simultano modelira spremenljivke na različnih nivojih brez uporabe agregacije ali deagregacije.

Modeli, ki se nanašajo na večnivojsko analizo, se v različni literaturi pojavljajo pod različnimi imeni:

- hierarhični linearni modeli (angl. hierarchical linear models) (Raudenbush in Bryk 2002),
- večnivojski regresijski modeli (angl. multilevel regression models) (Hox 2002),
- modeli komponent variance (angl. variance component model) (Longford v Hox 2002) oziroma modeli komponent kovariance (angl. covariance components models) (Roux 2006),
- modeli mešanih učinkov oziroma mešani modeli (angl. mixed effects or mixed model) (Littell, Milliken, Stroup in Wolfinger v Hox 2002),
- modeli naključnih koeficientov (angl. random coefficient model) (Leeuw, Kreft in Longford v Hox 2002) oziroma regresijski modeli naključnih koeficientov (angl. random-coefficient regression models) (Rosenberg in Longford v Raudenbus in Bryk 2002) oziroma modeli naključnih učinkov (angl. random effects models) (Roux 2006),
- modeli kontekstualnih učinkov (angl. contextual effects models) (Roux 2006),
- linearni mešani modeli (angl. linear mixed models) (Garson 2008).

V svoji nalogi bom zaradi bolj splošne indikacije te modele poimenovala večnivojski modeli (angl. multilevel models) (Snijders 1999, Bryk in Raudenbush 2002, Roux 2006, Goldstein 1995, Kreft 1998).

Z uporabo večnivojskega modeliranja lahko celotno varianco razdelimo na varianco znotraj skupin in varianco med skupinami, zato se ločijo tudi modeli struktur variance. Zmožnost ocenitve komponent variance (ki zagotavlja pomembno informacijo glede variabilnosti v rezultatih med in znotraj skupin) je ena od ključnih značilnosti večnivojskih modelov in ravno ta lastnost razlikuje večnivojske modele od tradicionalnih. Posledično se večnivojski modeli velikokrat imenujejo tudi modeli komponent variance (angl. variance components models). Nekaj tipičnih modelov:

- Modeli naključnih presekov (angl. Random intercept models)

Najenostavnejši večnivojski modeli vsebujejo le eno vrsto reziduala na vsakem nivoju. To vpliva na razčlenitev rezidualne variance na dve komponenti: med šolami in znotraj šol; zaradi tega se tak model pogosto imenuje tudi model komponent variance (angl. variance components models). Model se imenuje tudi model naključnih koeficientov, saj v regresijski enačbi le pojmu preseka dopuščamo, da naključno variira čez enote višjega nivoja (npr. šole). Za učinke pojasnjevalnih spremenljivk namreč predpostavljamo, da so enaki za vsako enoto višjega nivoja (šole). Slikovni prikaz bi torej predstavljale vzporedne črte (Rasbash, Steele, Browne in Prosser 2004).

- Model naključnih nagibov (angl. random slope models)

Pogosto nerealistična predpostavka modelov naključnih presekov je konstantnost učinkov pojasnjevalnih spremenljivk čez enote višjega nivoja. V primeru izobraževanja, denimo, se lahko šole razlikujejo glede učinka na dosežke pri starosti učencev 16 let (Y) glede na njihovo nadarjenost (X). V tem primeru slikovnega prikaza ne bi predstavljale vzporedne linije, saj bi linije, ki predstavljajo povezavo med X in Y, imele različne nagibe. To dosežemo s podrobno označitvijo dveh rezidualov na višjem nivoju: rezidual preseka ter rezidual nagiba, ki lahko med seboj korelirata. Naključni nagib za X nam pove tudi, da je varianca med šolami odvisna od spremenljivke X. Takšne vrste modeli so z bolj splošnim izrazom poimenovani tudi modeli naključnih koeficientov (angl. random coefficient models) (Rasbash, Steele, Browne in Prosser 2004).

- Kompleksna variacija nivoja 1 – heteroscedastičnost (angl. complex level 1 variation – heteroscedasticity)

V modelu naključnih nagibov koeficienti ene ali več pojasnjevalnih spremenljivk (X) naključno variirajo čez enote višjega nivoja. Ta model predpostavlja, da je varianca med enotami višjega nivoja kvadratna funkcija X. Vendar pa je za varianco med enotami nižjega nivoja znotraj enot višjega nivoja (npr. med učenci znotraj šol) predpostavljeno, da je konstantna (homoscedastičnost). Ta predpostavka je lahko nesmiselna, saj lahko pogosto opazimo, da na primer fantje variirajo bolj kot punce pri učnem uspehu, kar pomeni, da gre za

heteroscedastičnost na nivoju 1. Torej v modelu kompleksne variance nivoja 1 dopuščamo, da je varianca odvisna od pojasnjevalne spremenljivke (Rasbash, Steele, Browne in Prosser 2004).

- Autokorelacija – modeli časovnih vrst (angl. autocorrelation – time series models)

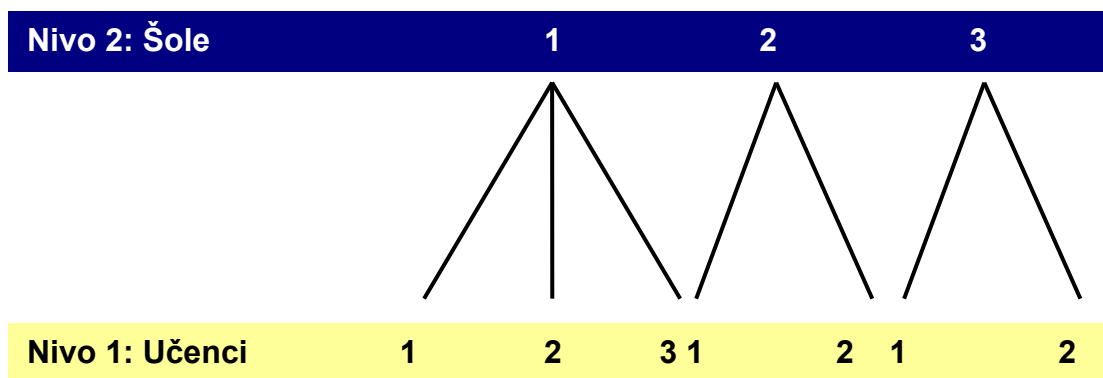
V standardnem dvonivojskem modelu predpostavljamo, da je korelacija parov med odgovori enot nižjega nivoja, naključno izbranih iz iste enote višjega nivoja, enaka ne glede na to, kateri par izberemo. Ta predpostavka konvertibilnosti pa je lahko vprašljiva za podatke s ponavljajočimi merjenji. Kot primer – pričakujemo lahko, da bo korelacija med odgovori razpadla kot funkcija časovnega razmika med merjenji, kar nas pripelje do avtokorelacijskih rezidualov na prvem nivoju. Avtokorelacijo lahko vključimo z uporabo večnivojskega modela časovnih vrst (angl. multilevel time series models) (Rasbash, Steele, Browne in Prosser 2004).

3. PODATKOVNE STRUKTURE

Veliko vrst podatkov je strukturiranih hierarhično, saj ima tudi večina raziskovalnih vprašanj v resničnem svetu kompleksno strukturo. Hierarhijo v večnivojski analizi predstavljajo enote, ki so združene znotraj različnih skupin. Najdemo jo na vseh področjih raziskovanja v šolstvu, kjer hierarhijo predstavljajo, denimo, šole in učenci znotraj šol, v geografiji, kjer hierarhijo predstavljajo države in kraji znotraj njih, v psihologiji družine in posamezniki znotraj njih in še bi lahko naštevali. Tako je v določenem smislu osnovni problem sociologije najti povezavo med lastnostmi posameznikov in lastnostmi skupin in struktur, v katerih posamezniki delujejo. V primeru, ko imamo učence združene znotraj razredov, tako pridobimo spremenljivke, ki opisujejo učence in spremenljivke, ki opisujejo razrede. Pri vzpostavitvi hierarhične strukture lahko dobimo agregirane spremenljivke učencev (število učencev v razredu, povprečni družbeno ekonomski status), ki opisujejo razrede. Še več, z agregiranimi spremenljivkami razredov in učencev lahko opišemo tudi učitelje, ki jih poučujejo, šole, v katerih se nahajajo, in soseske, v katerih se te šole nahajajo.

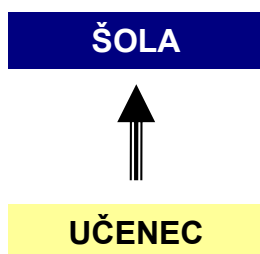
Kompleksnost podatkov lahko razširimo na več hierarhičnih nivojev, če na primer v šolstvu zraven šol in učencev uporabljamo še razrede, ali v geografiji, če poleg držav in mest dodamo še pokrajine. S tem lahko potrdim tezo Krefta, ki je trdil »Ko enkrat ugotoviš, da hierarhije obstajajo, jih zaznaš vsepovsod« (Kreft v Jones 2006: 11). V večnivojski analizi se te hierarhije nanašajo na nivoje. V šolstvu (glej sliko 3.1) bi torej prvi nivo predstavljali učenci in drugi nivo šole, v geografiji (glej sliko 3.3) pa prvi nivo predstavljajo mesta (npr. Maribor, Barcelona,...), drugi nivo pokrajine (Štajerska, Katalonija,...) in tretji nivo države (Slovenija, Španija,...).

Slika 3.1: DIAGRAM ENOT DVONIVOJSKE HIERARHIJE



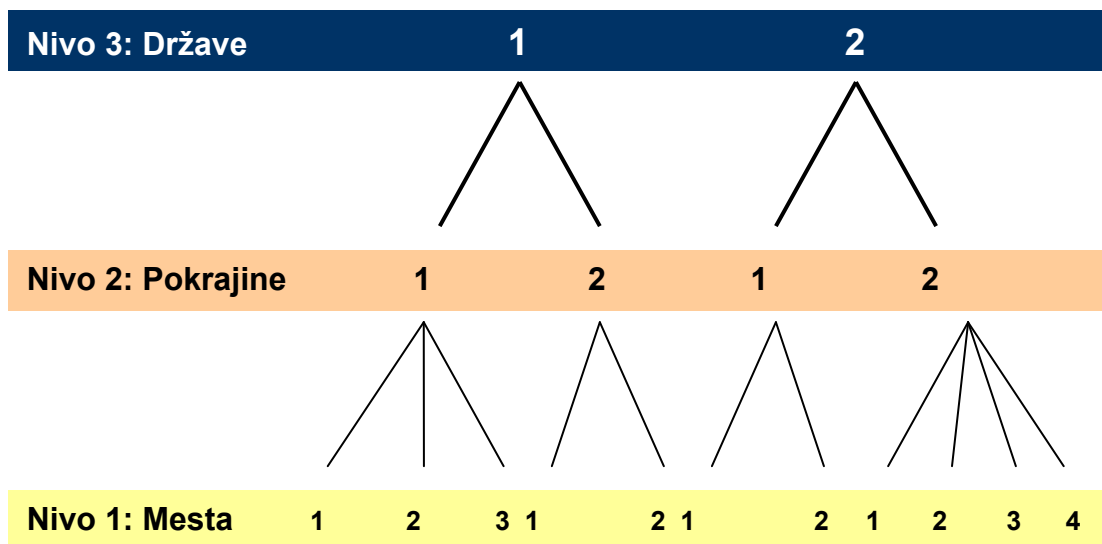
Vir: Center for Multilevel Modelling 2008.

Slika 3.2: KLASIFIKACIJSKI DIAGRAM DVONIVOJSKE HIERARHIJE



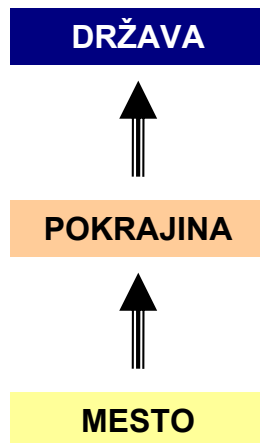
Vir: Center for Multilevel Modelling 2008.

Slika 3.3: DIAGRAM TRINIVOJSKE HIERARHIJE



Vir: Center for Multilevel Modelling 2008.

Slika 3.4: KLASIFIKACIJSKI DIAGRAM TRINIVOJSKE HIERARHIJE

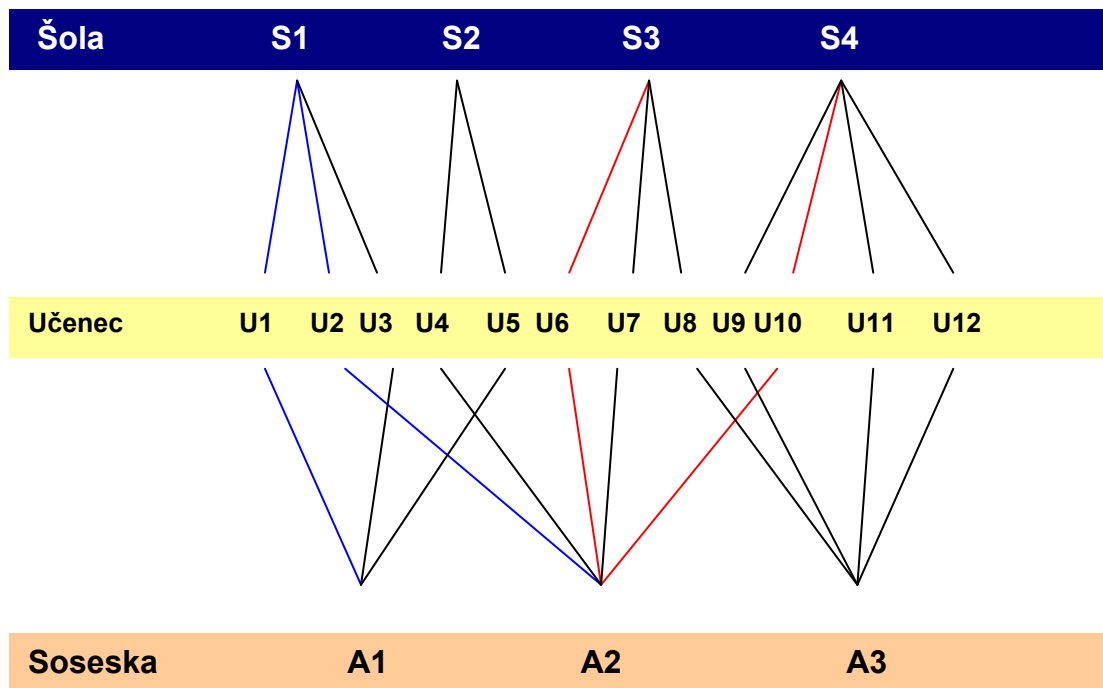


Vir: Center for Multilevel Modelling 2008.

Specifičnost takšnih podatkov, ki se združujejo glede na enote višjega nivoja, je njihova podobnost. Takšne enote so si vsekakor veliko bolj podobne, kot če bi te enote pridobili z naključnim izborom iz populacije. Dva posameznika, ki izhajata iz iste družine, imata veliko več skupnih lastnosti kot dva posameznika, ki izhajata iz dveh različnih družin. Podobnosti, ki jih v tem primeru lahko opazimo, so denimo fizična podobnost, uspeh v šoli in bolezn, kar kaže na to, da so te podobnosti lahko tako fizične kot tudi psihične. Seveda pa se glede na način združevanja enot v višje nivoje ta podobnost karakteristik spreminja. V šolah, kjer so učenci naključno razdeljeni v razrede, ponavadi obstaja manj fizičnih podobnosti med njimi, obstajajo pa na primer podobnosti glede starosti, učiteljev, ki jih poučujejo, ali krajev, v katerih živijo.

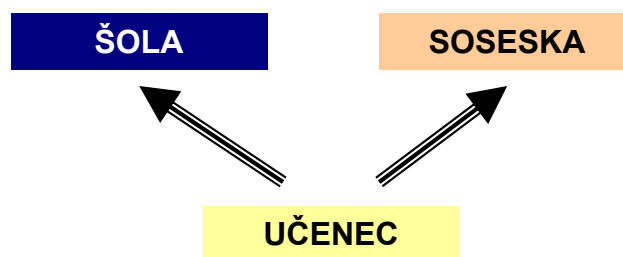
Večnivojsko modeliranje zavzema tudi modele, kjer enote niso združene le na enem nivoju, ampak so prepletene. Eden takšnih primerov je dvojna navzkrižna razvrstitev (angl. two-way cross-classification).

Slika 3.5: DIAGRAM NAVZKRIŽNE KLASIFIKACIJE



Vir: Center for Multilevel Modelling 2008.

Slika 3.6: KLASIFIKACIJSKI DIAGRAM NAVZKRIŽNE KLASIFIKACIJE



Vir: Center for Multilevel Modelling 2008.

Slika 3.5 prikazuje diagram enot, kjer se učenci nahajajo znotraj navzkrižne razvrstitve šol glede na soseske. V tem primeru šole niso združene znotraj sosesk, zato tudi učenci ne morejo biti združeni znotraj šol in sosesk hkrati. Učenca U1 in U2 obiskujeta isto šolo, vendar prihajata iz različnih sosesk, učenca U6 in U10 pa prihajata iz iste soseske, a ne obiskujeta iste šole. Podoben primer takšnega modela je model multiplega članstva (angl. multiple membership models), kjer lahko enote nižjega nivoja sočasno pripadajo več enotam višjega nivoja.

4. PREDNOSTI VEČNIVOJSKE ANALIZE

V primeru, ko razpolagamo s podatki na multiplih nivojih, le-ti vsebujejo vplive skupine na posameznike, kar je lahko v primeru tradicionalnih statističnih tehnik ocenjeno nepravilno oziroma so ti vplivi neupoštevani. Ko imamo prisotne razvrstitve (na primer študentje na fakultetah, zaposleni v organizacijah,...), so opazovanja znotraj skupine pogosto med seboj bolj podobna, kot bi to lahko zaznali v primeru podatkov brez razvrstitve v skupine. V takšnem primeru regresija prezre učinke skupin, kar lahko omaja natančnost sklepanja. Tradicionalne statistične tehnike enote analize obravnavajo kot neodvisna opazovanja. Posledica nezaznavanja hierarhičnih struktur je podcenjenost standardnih napak in regresijskih koeficientov, ki pelje v preveliko poudarjanje statistične signifikance. Standardne napake za koeficiente neodvisnih spremenljivk višjega nivoja so zaradi neupoštevanja skupin popačene. Bistveni pomen vplivov skupin se izraža v večini situacij, kjer je ključno raziskovalno vprašanje obseg grupiranja izidov posameznikov in identifikacija skupin, v katerih se nahajajo. Skupine, v katerih se posamezniki nahajajo, imajo lahko pomembne vplive na posameznikovo delovanje, zato je takšne vplive potrebno upoštevati. Če ponovno vzamemo primer šol, se pri vrednotenju kvalitete šole lahko osredotočimo na ocenjevanje vplivov šol na dosežke učencev, ki se znotraj njih nahajajo. Bolj kvalitetne šole imajo verjetno bolj stimulative vplive na dosežke svojih učencev kot manj kvalitetne šole. Če želimo pri tradicionalnem regresijskem modelu (vsota najmanjših kvadratov) upoštevati vplive skupin, moramo vanj vključiti sintetično oziroma umetno spremenljivko (angl. dummy variable). Takšen model se imenuje analiza variance ali model fiksnih vplivov. V veliko primerih pa bodo neodvisne spremenljivke definirane na skupinskem nivoju, kot je na primer tip šole – mešane šole ali šole ločene glede na spol. V modelu fiksnih vplivov bodo vplivi neodvisnih spremenljivk skupinskega nivoja pomešani z vplivi skupinskih sintetičnih/umetnih spremenljivk, kar pomeni, da ni mogoče ločiti vplivov glede na opazovane in neopazovane lastnosti skupin. V večnivojskem modelu (model naključnih vplivov) lahko posebej ocenimo vplive obeh vrst spremenljivk. Pri modelu fiksnih vplivov se pojavi tudi problem sklepanja na populacijo. V večnivojskem modelu

so namreč skupine v vzorcu obravnavane kot naključni vzorec iz populacije skupin. V primeru modela fiksnih vplivov je nemogoče sklepati na celotno populacijo skupin oziroma na skupine, ki niso bile vključene v analizo (Rasbash 2006).

4.1 TRADICIONALNI PRISTOPI REGRESIJSKEGA MODELIRANJA VEČNIVOJSKIH PODATKOV

Obstajajo trije tradicionalni pristopi regresijskega modeliranja večnivojskih podatkov:

ENOSTAVNA REGRESIJA, imenovana tudi naivna regresija, ki prezre učinke višjih nivojev. Tak pristop je primeren le v primeru, ko je raziskovalec prepričan, da učinki višjih nivojev ne obstajajo. Vendar v večini primerov takšni učinki obstajajo, kar pomeni, da enostavna regresija vodi do podcenjevanja standardnih napak, precenjevanja napak prve vrste in do preozke mere zaupanja v primerjavi z večnivojskim modeliranjem istih podatkov.

REGRESIJA FIKSNIH UČINKOV oziroma deagregacija. Priljubljen tradicionalni pristop, s katerim deagregiramo podatke na nižji nivo (tipičen primer bi bil, da so vsakemu študentu pripisane različne spremenljivke nivoja šole in za enote analize uporabimo le študente), kar pomeni, da enotam na nižjem nivoju pripišemo spremenljivke višjega nivoja. Torej imajo enote iz iste skupine iste vrednosti na dobljenih kontekstualnih spremenljivkah. Problem nastane, kadar vemo, da študentje spadajo v isto šolo in tako vemo tudi, da imajo ti elementi pri spremenljivkah šol isto vrednost. Potemtakem ne moremo predpostavljati neodvisnosti opazovanja, ki je osnova klasičnih statističnih tehnik (Raudenbush in Bryk 2002). To pa ni edini problem, s katerim se soočamo. Tipična je tudi situacija, ko so upoštevane le napake vzorčenja za prvi nivo, medtem ko so le-te za drugi nivo neupoštevane. To pomeni, da so informacije enot višjih nivojev obravnavane nepravilno, kot da bi bile neodvisni podatki za enote nižjega nivoja. Neupoštevanje teh vrednosti pripelje do preveč optimističnih ocen signifikanc.

Nenazadnje je pomembno tudi dejstvo, da v primeru regresije fiksnih učinkov število umetnih spremenljivk (angl. dummy variables) narašča z naraščanjem števila skupin, kar vodi v neučinkovitost ocen. V večnivojskem modeliranju pa je dodajanje navzkrižno nivojskih interakcij možno in le-te doprinesejo k izboljšanju rezultatov. Model namreč ločeno upošteva preseke in nagibe za posameznike v vsaki skupini višjega nivoja, kjer je spremenljivka skupine upoštevana kot naključni učinek.

REGRESIJA POVPREČIJ MERJENJ oziroma agregacija (angl: summary measures regression). Še en tradicionalni pristop k večnivojskim problemom, kjer agregiramo podatke na višji nivo (primer: šolam pripišemo povprečja rezultatov učencev in za enote analize uporabimo šole). Tako pridobljeni podatki torej predstavljajo povprečja in so ponavadi centrirani okrog ničle. »Največji problem, ki se nam pojavi pri tovrstni analizi, je, da še preden začnemo z analizo, izgubimo pomembnejše znotrajskupinske informacije, ki lahko predstavljajo tudi do 80 ali 90 odstotkov celotne variacije« (Raudenbush in Bryk 2002). V ta namen uporabimo običajno regresijo najmanjših kvadratov (OLS regression). Problem z regresijo povprečij merjenj je podagregacija, saj manjše število enot analize na višjem nivoju nadomesti večje število enot na nižjem nivoju, kar se odraža v izgubi statistične moči. Tako kot v primeru enostavne regresije, opisane pod prvo točko, tudi ta vrsta analize vodi v podcenjevanje standardnih napak, precenjevanje vrednosti napak prve vrste ter podaja preozke mere zaupanja.

Konceptualni problem, ki se nanaša na to vrsto analize, je nepazljivost pri interpretaciji rezultatov, saj lahko pripelje do zmote na napačnem nivoju. Ta je seveda lahko posledica analiziranja enot na enem nivoju in podajanje sklepov na drugem nivoju. Najbolj poznana napaka je ekološka zmota¹ (angl. ecological fallacy), pri kateri gre za interpretacijo agregiranih podatkov na nivoju posameznika, torej se povezave na višjem nivoju pripisujejo povezavam na nižjem nivoju. Robinson podaja agregirane podatke, ki opisujejo povezavo med deležem temnopoltih in nivojem nepismenosti v devetih zveznih državah

¹ Ali Robinsonov učinek po avtorju (angl. *Robinson Effect*) (Hox 2002).

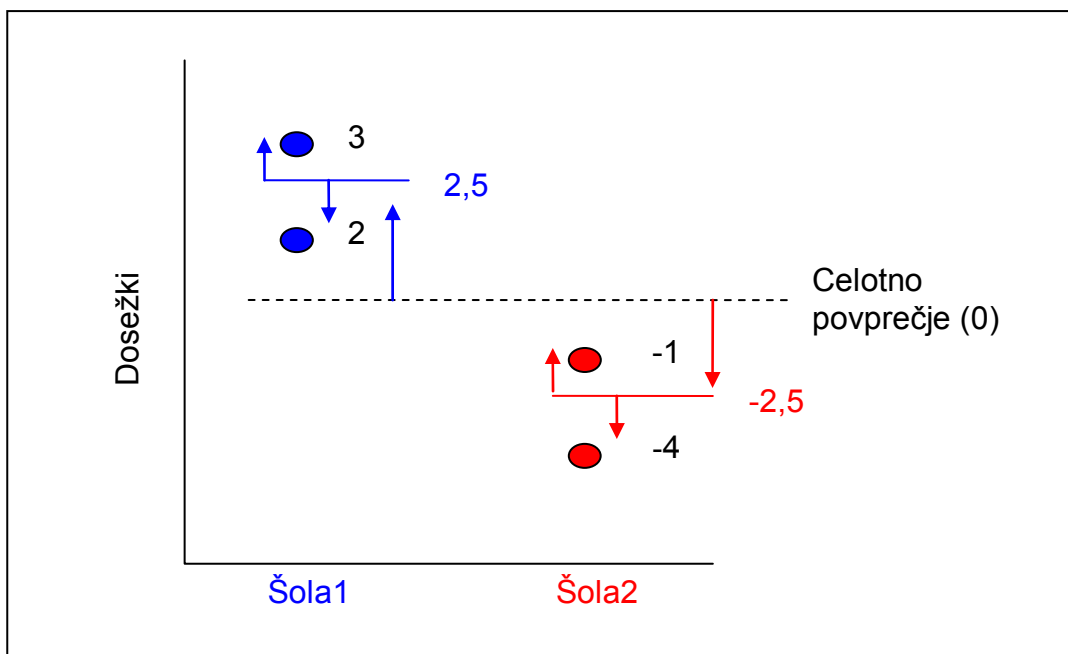
Združenih Držav Amerike v letu 1930. Ekološka korelacija oziroma korelacija med agregiranimi spremenljivkami na nivoju pokrajin je znašala 0,95. Na drugi strani je korelacija nivoja posameznikov med temi globalnimi spremenljivkami znašala 0,20. Robinson je zaključil, da v praksi ekološka korelacija zagotovo ni enaka ustrezni korelaciji nivoja posameznikov. V Združenih Državah Amerike sta rasa in izobraženost sicer v majhnem soodnosu na nivoju posameznikov, a je ta soodvisnot nekoliko višja na nivoju držav, saj imajo južnejše države veliko več Afroameričanov in tudi veliko višjo nepismenost pri vseh rasah. Takšen problem se lahko pojavi tudi v obratni smeri, ko podajamo sklepe o višjem nivoju na podlagi analize, izvedene na nižjem nivoju. Takšna vrsta zmote se imenuje atomistična zmeta (angl. atomistic fallacy). Povezana, a nekoliko drugačna zmeta, je poznana kot Simpsonov paradoks (angl. Simpson's Paradox). Ta se nanaša na problem, da lahko pridobimo popolnoma napačne ugotovitve, če so grupirani podatki, ki izhajajo iz heterogene populacije, analizirani, kot da podatki izhajajo iz ene homogene populacije (Hox 2002).

4.2 PRAKTIČEN PRIMER – UPORABNOST VEČNIVOJSKEGA MODELA

Osnovno vprašanje, ki bi ga rada pojasnila s praktičnim primerom, se glasi: *koliko variabilnosti v učnih dosežkih lahko pripišemo faktorjem nivoja šole in koliko faktorjem nivoja učencev?*

Za ponazoritev tega posplošenega primera bom uporabila dve šoli, vsako z dvema učencema, kar pomeni, da ima ta model dva nivoja. Prvi nivo predstavljajo učenci, drugega pa šole. V šoli 1 je uspeh za prvega učenca 3 in za drugega 2. V šoli 2 učenca dosegata uspeh -1 in -4. Diagram enot dvonivojskega združevanja za primer šol in učencev prikazuje Slika 4.1.

Slika 4.1: GRAFIČNI PRIKAZ ENOSTAVNEGA PRIMERA



Vir: Rasbash 2006.

Celotno povprečje učnih dosežkov je 0^2 in celotna varianca je $7,5^3$. Glede na to, da imamo enote razvrščene v skupine (učence glede na šole), lahko izračunamo variance povprečja šol okrog celotnega povprečja. Povprečje šole 1 je tako 2,5, šole 2 pa -2,5. Varianca povprečja šol okrog celotnega povprečja je $6,25^4$, izračunamo pa lahko tudi varianco učnih rezultatov učencev okrog

² $(3+2+(-1)+(-4))/4=0$

³ $(9+4+1+16)/4=7,5$

⁴ $((2,5)^2 + (-2,5)^2)/2=6,25$

povprečja šol, ki znaša $1,25^5$. Skupna varianca je potemtakem tudi seštevek obeh izračunanih varianc; $1,25 + 6,25 = 7,5$.

S tem primerom sem operacionalizirala variabilnost, kar mi dovoli prikaz celotne variance in razlago, koliko te variance pojasnjujejo šole (6,25 od 7,5) in koliko učenci (1,25 od 7,5). Glede na izračunano ugotovimo, da 82 odstotkov⁶ celotne variance v učnih dosežkih lahko pripišemo faktorjem drugega nivoja – šolam in 18 odstotkov⁷ celotne variance faktorjem prvega nivoja – učencem. Pri tem je potrebno poudariti, da ta primer le nakazuje prednosti večnivojskega modeliranja pri analizi variance, zato so tudi rezultati zgolj informativni. V večini resničnih družbenih sistemov je najbolj variabilen nivo ravno tisti, kjer se nahajajo ljudje (Rasbash 2006).

Seveda je glede na vrste podatkov smotno preučiti, katera vrsta analize je najbolj primerna, da dobljeni rezultati kar najbolj predstavljajo realnost. Tako obstaja veliko raziskovalnih problemov, ki jih je verjetno celo bolje modelirati kot osnovne linearne modele, a v primeru, ko imamo podatke z združenim virom variabilnosti, je vsekakor bolj primerna večnivojska analiza. Če povzamemo, večnivojska analiza zagotavlja naslednje prednosti:

- običajni linearni modeli predpostavljajo neodvisnost opazovanj, medtem ko je večnivojska analiza ne,
- običajni linearni model ponavljajočih merjenj predpostavlja, da so bile vse meritve na subjektih opravljene v isti časovni točki, medtem ko večnivojska analiza dopušča merjenje na različnih časovnih točkah,
- v primeru ponavljajočih merjenj običajni linearni modeli zahtevajo enako število ponovljenih merjenj, medtem ko večnivojska analiza dopušča neenaka ponavljanja. To pomeni, da je večnivojska analiza enako učinkovita za uravnotežene in neuravnotežene načrte, medtem ko so običajni linearni modeli učinkoviti le v primeru uravnoteženih načrtov,

⁵ $((3-2,5)^2 + (2-2,5)^2 + (-1-(2,5))^2 + (-4-(-2,5))^2) / 4 = 1,25$

⁶ $6,25 / 7,5$

⁷ $1,25 / 7,5$

- običajni linearni modeli zahtevajo, da so vse interakcije faktorjev znotraj subjektov in med subjekti vključene v analizo, medtem ko večnivojska analiza dopušča, da raziskovalec vključi interakcije glede na interese raziskave,
- običajna linearna analiza podaja točno določene predpostavke glede kovariančnih matric, medtem ko pri večnivojski analizi to ni tako natančno in dopušča več možnosti,
- večnivojska analiza podpira podatke s hierarhično strukturo (podatke na nižjem nivoju združene v enote na višjem nivoju), medtem ko običajni linearni modeli tega ne podpirajo,
- ocene večnivojske analize temeljijo na metodi maksimalne verjetnosti (angl. maximum likelihood – ML) in metodi omejene maksimalne verjetnosti (angl. restricted maximum likelihood – REML), medtem ko običajna linearna analiza temelji na metodi ANOVA (Garson 2005).

5. PREDPOSTAVKE REGRESIJE IN VEČNIVOJSKE ANALIZE

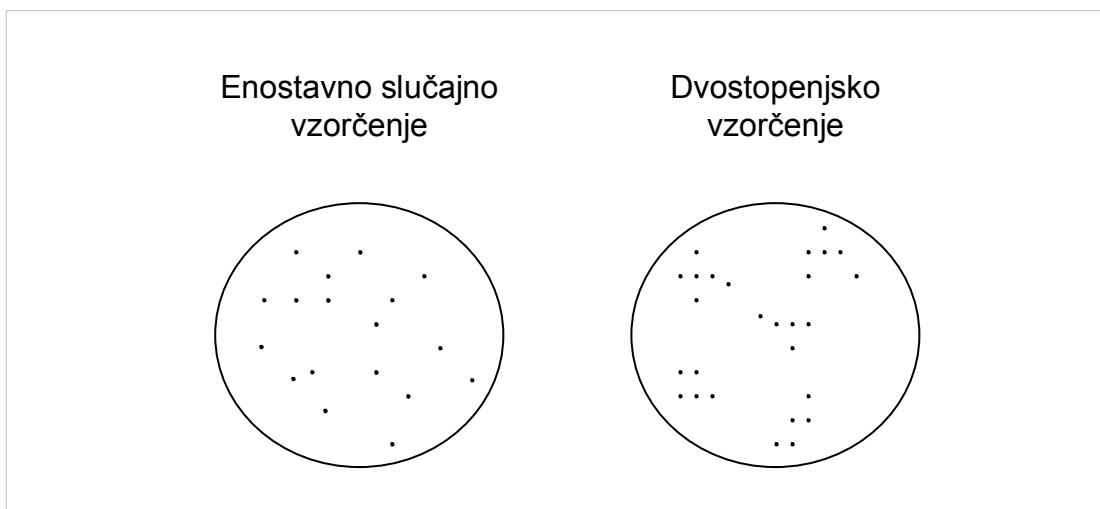
Če se omejimo le na tradicionalne linearne modele analize, opazimo, da so osnovne predpostavke te analize linearnost (angl. linearity), normalnost porazdelitve rezidualov (angl. normality), homoscedastičnost (angl. homoscedasticity) in neodvisnost opazovanj (angl. independence). V večnivojski analizi poskušamo ohraniti prvi dve predpostavki, medtem ko moramo zadnji dve, predvsem predpostavko o neodvisnosti, prilagoditi. Osnovna ideja te prilagoditve je dejstvo, da so si posamezniki v isti skupini bližje in med seboj bolj podobni. Kljub temu, da so študentje v različnih šolah neodvisni, pa si le-ti delijo iste vrednosti na veliko več spremenljivkah. Nekaj takšnih spremenljivk ne bo merjenih, kar pomeni, da se bodo izgubile v intervalu napake linearnega modela, kar povzroči korelacijo med motnjami. To idejo lahko natančneje razvijemo z uporabo modela komponent variance (angl. variance components models). Te motnje namreč vsebujejo elemente posameznikov in skupin. Elementi posameznikov so neodvisni, medtem ko so elementi skupin neodvisni med skupinami, a popolnoma korelirajo znotraj skupin. Nekateri skupine so lahko bolj homogene kot druge, kar pomeni, da se varianca elementov skupin lahko razlikuje (Raudenbush in Bryk 2002).

5.1 VZORČENJE

V regresiji je ena ključnih predpostavk uporaba enostavnega naključnega vzorca (angl. SRS – Simple random sample). Vendar pa so enote, ki jih potrebujemo za večnivojsko analizo, specifične glede na svojo pripadnost enotam višjega nivoja, zato v tem primeru enostavno slučajno vzorčenje ni primerno. Pri večnivojski analizi tako uporabimo dvostopenjsko vzorčenje (angl. two stage sampling) oziroma dvostopenjsko vzorčenje v skupinah (angl. two stage cluster sampling). V tem primeru uporabljamo že omenjeno hierarhijo, kjer najprej izberemo enote višjega nivoja oziroma večje skupine, nato enote nižjega nivoja, dokler ne pridemo do nivoja, kjer izbiramo elemente. Pri raziskavi med učenci bi tako najprej izbrali vzorec šol, nato vzorec razredov in nazadnje vzorec učencev. Zato se v splošnem tak način vzorčenja imenuje tudi

večstopenjsko vzorčenje (angl. multi-stage sampling), saj zaporedno vzorčimo na več elementih.

Slika 5.1: RAZLIKA MED ENOSTAVNIM SLUČAJNIM VZORČENJEM IN DVOSTOPENJSKIM VZORČENJEM



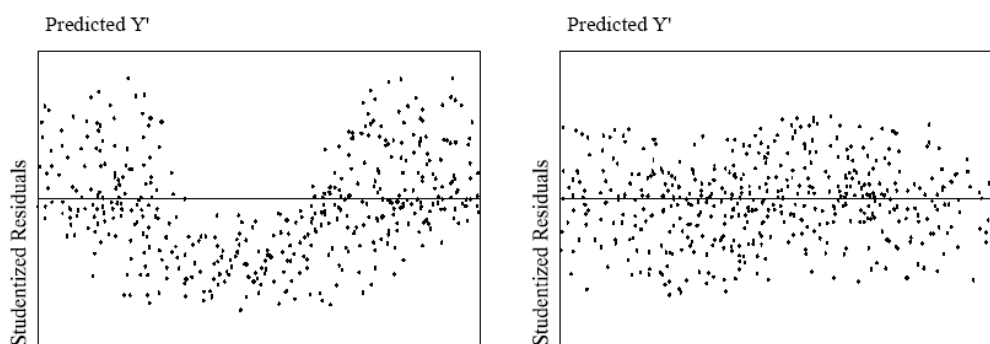
5.2 NEODVISNOST OPAZOVANJ

Pri večnivojski analizi se uporabljajo populacije s hierarhično strukturo in, kot že rečeno, se ta način vzorčenja imenuje večstopenjsko vzorčenje. Pri tem najprej izberemo enote višjega nivoja (šole) in nato znotraj teh enot izberemo še elemente nižjega nivoja (učence). V takšnih primerih zato posamezna opazovanja niso popolnoma neodvisna. V primeru šol in učencev se to pokaže v večji podobnosti učencev znotraj iste šole. Gre torej za situacijo, ko so odvisne spremenljivke pri opazovanju na nižjem nivoju združene znotraj istih enot višjega nivoja povezane. Rezidualna korelacija tako krši predpostavko o neodvisnosti opazovanj, ki je osnovna predpostavka običajnih regresijskih pristopov. Večnivojski modeli torej obravnavajo podatke, kjer so opazovanja neneodvisna, kar pomeni, da pravilno modelirajo soodnosne napake, medtem ko jih običajna linearna regresija ne.

5.3 LINEARNOST

Linearnost predpostavlja povezavo med odvisnimi in neodvisnimi spremenljivkami kot ravno oziroma linearno premico. To pomeni, da sprememba spremenljivke X sproži sorazmerno spremembo spremenljivke Y. V primeru, da povezava med odvisno in neodvisno spremenljivko ni linearna, rezultati, ki jih pridobimo, podcenjujejo resnično povezavo med njima.

Slika 5.2: PRIMER KRIVULJASTE LINEARNE IN LINEARNE POVEZAVE



Vir: Osborne in Waters 2002: 2.

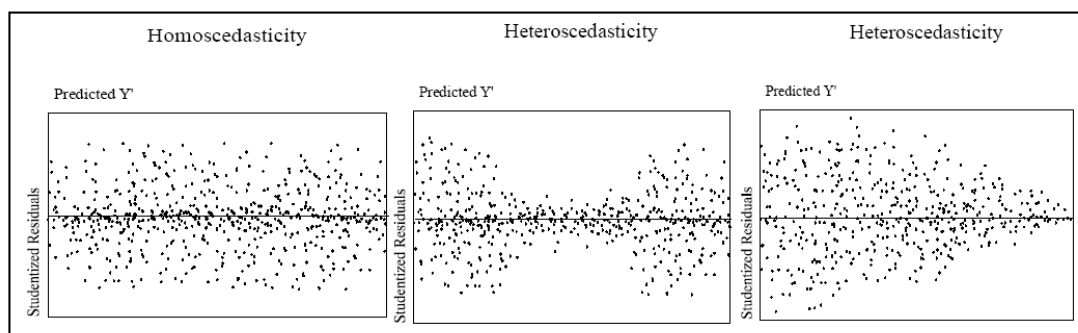
5.4 HOMOSCEDASTIČNOST

Naslednja izmed predpostavk linearne regresije je homoscedastičnost in se nanaša na podajanje veljavnih statističnih povzetkov o povezavah v populaciji. Pri tem gre za to, da sta standardni odklon in varianca v smislu napake (μ) konstantna za vse neodvisne spremenljivke in da pojmi napak izhajajo iz iste populacije oziroma je varianca napake enaka za različne vrednosti X. To nam pove tudi, da obstaja enoten graf za prikaz regresijske premice, kjer so enote združene v elipsasto obliko. Če predpostavka ne drži, naši rezultati niso veljavni. To torej pomeni, da so vse napake neodvisne od pojasnjevalnih spremenljivk in da je varianca okrog regresijske premice enaka za vse neodvisne spremenljivke ter od nje ni odvisna.

V nasprotju s homoscedastičnostjo pojem heteroscedastičnost predstavlja situacijo, ko variance odvisnih spremenljivk variirajo med podatki. Heteroscedastičnost pomeni, da variance (razpršitev okrog pričakovane sredine

ničle) rezidualov niso konstantne, ampak različne za različna opazovanja. To lahko v regresijski analizi predstavlja problem, saj je v primeru neenakih varianc tudi relativna zanesljivost za posamična opazovanja neenaka. Večja kot je namreč varianca, manjše uteži bi morale biti pripete na opazovanja. Problem je bolj očiten, ko je vrednost variance povezana z eno ali več neodvisnimi spremenljivkami. Porazdelitev rezidualov namreč ne bi smela imeti povezave z nobeno od spremenljivk, kar je predpostavka klasičnega modela. Homoscedastičnost je potemtakem pri večnivojski analizi rahlo iznakažena, saj so v različnih skupinah variance različne zaradi različnosti skupin višjega nivoja.

Slika 5.3: GRAFIČNI PRIKAZ HOMOSCEDASTIČNOSTI IN HETEROSCEDASTIČNOSTI



Vir: Osborne in Waters 2002: 4.

5.5 NORMALNOST PORAZDELITVE REZIDUALOV

Običajna regresijska analiza najmanjših kvadratov (OLS) predpostavlja, da so reziduali⁸ normalno porazdeljeni in neodvisni s povprečjem nič in s konstantno varianco. Ko podatke zbiramo s skupinsko metodo vzorčenja, reziduali niso neodvisni drug od drugega. Reziduali so v primeru neodvisne porazdelitve naključno razpršeni okrog ničle. V primeru, ko reziduali niso tako enakomerno porazdeljeni okrog premice, gre za heteroscedastičnost, ki je razložena v prejšnjem podpoglavju.

⁸ Razlika med opazovanimi in pričakovanimi vrednostmi modela.

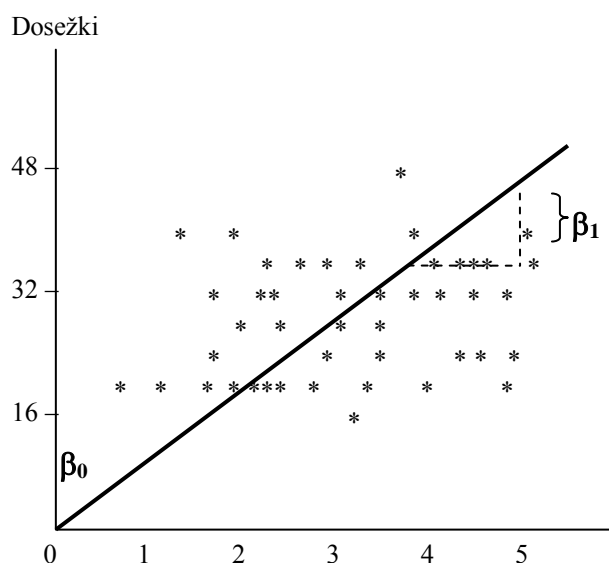
6. STATISTIČNO OZADJE VEČNIVOJSKE ANALIZE

V tem poglavju želim predstaviti statistično ozadje večnivojske analize. Princip je verjetno še najlažje razložiti s pomočjo enostavnega primera, ki temelji na razumevanju samega jedra ideje regresije in analize variance. Glede na to, da sem v večini literature zasledila uporabo večnivojske analize na področju izobraževanja oziroma šolstva, sem se odločila, da bom tudi sama predstavila primer iz tega področja.

6.1 RAZISKAVA POVEZAVE MED DRUŽBENO EKONOMSKIM STATUSOM IN USPEHOM V ŠOLI

Predpostavljajmo hipotetični odnos med eno neodvisno spremenljivko, ki se nanaša na študenta (denimo družbeno ekonomski status – DES) in eno samo odvisno spremenljivko (matematični uspeh) znotraj ene šole. Gre torej za osnovni model z eno odvisno spremenljivko na enem nivoju. Na sliki 6.1. je ta odnos prikazan na razsevnem grafikonu. Množico točk najbolje predstavlja ravna črta s presekom β_0 in nagibom β_1 .

Slika 6.1: POVEZAVA MED USPEHOM IN DRUŽBENO EKONOMSKIM STATUSOM V ENI HIPOTETIČNI ŠOLI



Vir: Raudenbush in Bryk 2002: 17.

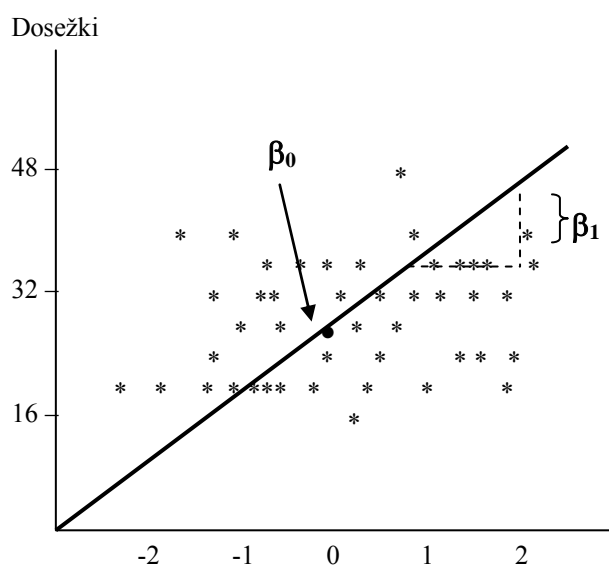
Regresijska enačba:

$$Y_i = \beta_0 + \beta_1 X_i + r_i \quad [\text{enačba 6.1}]$$

Presek β_0 je torej določen kot pričakovani matematični uspeh učenca, katerega DES je enak 0. Po drugi strani pa nagib β_1 predstavlja predvideno spremembo v matematičnem uspehu, ki jo sproži sprememba DES, če se le-ta poviša za eno enoto. Rezidual r_i (delež napake) predstavlja neodvisen učinek oziroma vpliv osebe i . Pri tem predpostavljamo, da je za r_i značilna normalna distribucija s povprečjem 0 in varianco δ^2 , tako da je $r_i \sim N(0, \delta^2)$.

Ponavadi je na sliki veliko bolj priročno prikazati neodvisno spremenljivko X tako, da ima presek določen pomen. Tako lahko denimo centriramo DES tako, da od vsake vrednosti odštejemo povprečni DES: $X_i - \mu_X$, kjer je μ_X , povprečni DES za to hipotetično šolo. Če sedaj prikažemo Y_i kot funkcijo $X_i - \mu_X$ (glej sliko 6.2) z regresijsko premico, lahko vidimo, da presek β_0 predstavlja povprečni matematični uspeh, medtem ko nagib ostaja nespremenjen.

Slika 6.2: POVEZAVA MED USPEHOM IN CENTRIRANO SPREMENLJIVKO DRUŽBENO EKONOMSKI STATUS V HIPOTETIČNI ŠOLI

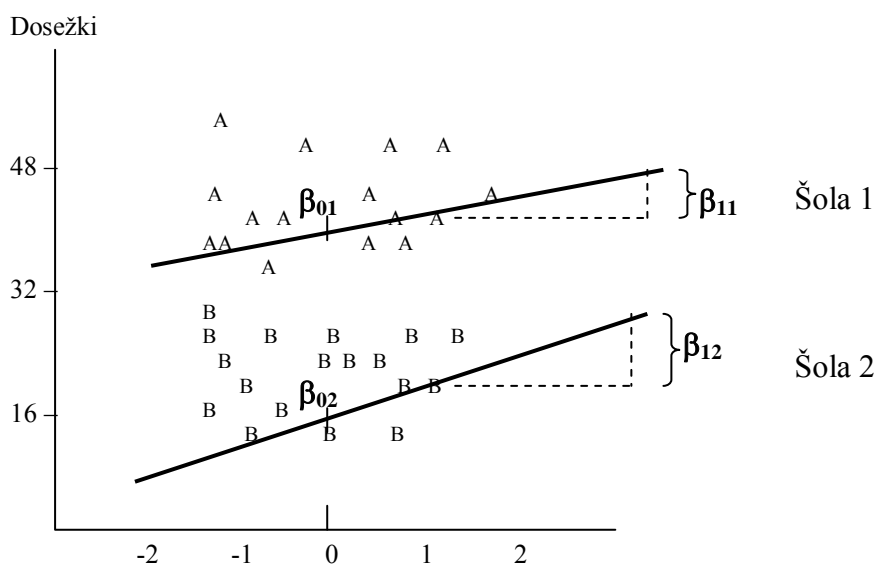


Vir: Raudenbush in Bryk 2002: 17.

6.2 RAZISKAVA POVEZAVE MED DRUŽBENO EKONOMSKIM STATUSOM IN USPEHOM V DVEH ŠOLAH

Sedaj ko smo predstavili sam koncept regresije, si poskušajmo predstavljati dve posamični regresiji za dve hipotetični šoli, ki sta predstavljeni na sliki 6.3. Položaj premic kaže, da se Šola 1 in Šola 2 razlikujeta v dveh ozirih. Prvič, Šola 1 ima višje povprečje kot Šola 2. Ta razlika se odraža v presekih obeh premic, tako da je $\beta_{01} > \beta_{02}$. Drugič, DES ima manjšo pojasnjevalno moč za matematični uspeh v Šoli 1, kot pa v primeru Šole 2, kar se odraža v primerjavi nagibov obeh premic, tako da je $\beta_{11} < \beta_{12}$.

Slika 6.3: POVEZAVA MED USPEHOM IN DRUŽBENO EKONOMSKIM STATUSOM ZNOTRAJ DVEH HIPOTETIČNIH ŠOL



Vir: Raudenbush in Bryk 2002: 18.

Če smo torej učence naključno porazdelili med obe šoli, lahko rečemo, da je Šola 1 tako bolj učinkovita v smislu matematične uspešnosti kot tudi bolj konstantna, kar pomeni, da je matematični uspeh bolj konstanten med učenci. Večja učinkovitost se kaže v višjem povprečju dosežkov v Šoli 1 ($\beta_{01} > \beta_{02}$). Večja enakomernost pa se odraža v manjšem nagibu ($\beta_{11} < \beta_{12}$). Seveda študenti ponavadi niso porazdeljeni po šolah naključno, zato ne moremo kar

tako sklepati o učinkih šole, ne da bi v ozir vzeli tudi razlike v sestavi študentov. Vendar pa ta predpostavka v tem primeru pomaga pojasniti cilje naše analize.

6.3 RAZISKAVA POVEZAVE MED DRUŽBENO EKONOMSKIM STATUSOM IN USPEHOM V J ŠOLAH

Predpostavljajmo, da imamo naključni vzorec J šol iz populacije in tako razširimo našo študijo odnosa med družbeno ekonomskim statusom in matematičnim uspehom na celotno populacijo šol. V tem primeru seveda ni več praktično prikazovati podatkov z razsevnim grafikonom za vsako šolo. Lahko pa opišemo opazovani odnos za vsako šolo j z naslednjo enačbo:

$$Y_{ij} = \beta_{0j} + \beta_{1j}(X_{ij} - \mu_{X.j}) + r_{ij}, \quad [\text{enačba 6.2}]$$

kjer spet poenostavljeno predpostavljamo, da je za r_{ij} značilna normalna porazdelitev in homogenost variance med šolami, torej $r_{ij} \sim N(0, \sigma^2)$. Preseku in nagibu je v tem primeru podpisan j , kar omogoča, da ima vsaka šola svoj edinstven presek in nagib. Učinkovitost in enakomernost porazdelitve vsake šole tako opredeljuje par vrednosti (β_{0j}, β_{1j}) . Tudi v tem primeru je pametno in priročno predpostavljati, da je za presek in nagib značilna bivariatna normalna distribucija znotraj populacije šol. Na ta način lahko opredelimo naslednje vrednosti:

$$E(\beta_{0j}) = \gamma_0, \quad \text{Var}(\beta_{0j}) = \tau_{00},$$

$$E(\beta_{1j}) = \gamma_1, \quad \text{Var}(\beta_{1j}) = \tau_{11},$$

$$\text{Cov}(\beta_{0j}, \beta_{1j}) = \tau_{01},$$

γ_0 je povprečna aritmetična sredina (preseki) šole v populaciji šol;

τ_{00} je populacijska varianca za aritmetične sredine (preseke) šol;

γ_1 je povprečni nagib odnosa DES-uspeh za populacijo;

τ_{11} je populacijska varianca nagibov;

τ_{01} je populacijska kovarianca med nagibi in preseki.

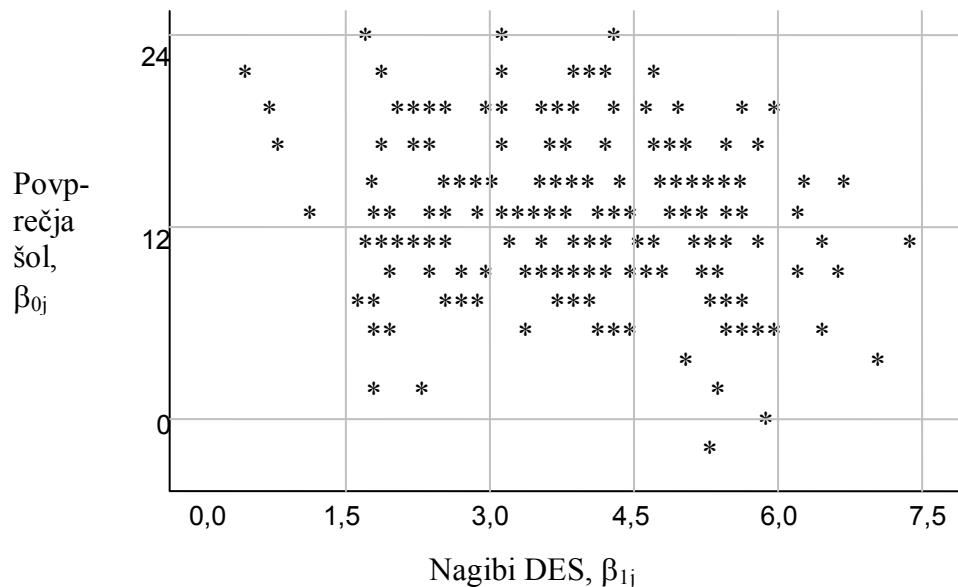
Pozitivna vrednost τ_{01} implicira, da imajo šole z visokim povprečjem (presekom) običajno tudi pozitivne nagibe. Poznavanje teh varianc in kovariance na ta način neposredno vodi do formule za izračun populacijske korelacije med povprečji in nagibi:

$$\rho(\beta_{0j}, \beta_{1j}) = \tau_{01} / (\tau_{00} \tau_{11})^{1/2} \quad [\text{enačba 6.3}]$$

Seveda v realnosti redko poznamo resnične vrednosti populacijskih parametrov, ki smo jih predstavili ($\gamma_0, \gamma_1, \tau_{11}, \tau_{00}, \tau_{01}$), kot tudi resnična povprečja in nagibe za vsako individualno šolo (β_{0j} in β_{1j}), ampak vse te parametre ocenimo na podlagi podatkov, a v tem poglavju želimo le prikazati pomene parametrov.

Če kljub temu predpostavljamo, da poznamo realne vrednosti povprečij in nagibov za vsako šolo, lahko narišemo graf 6.4, ki je nekakšen razsevni grafikon odnosov med β_{0j} in β_{1j} za hipotetični vzorec šol. Ta slika nam pove, kako šole med sabo variirajo v smislu svojih povprečij in nagibov. Tako lahko denimo opazimo večjo disperzijo povprečij (vertikalna os) kot nagibov (horizontalna os), kar lahko zapišemo kot $\tau_{00} > \tau_{11}$. Prav tako lahko opazimo, da oba učinka težita k negativni korelaciji. Šole z visokim povprečnim uspehom, β_{0j} , težijo k šibkemu odnosu DES-uspeh, β_{1j} , kar lahko zapišemo kot $\tau_{01} < 0$. Šole, ki so učinkovite in z enakomerno porazdelitvijo – imajo visok povprečni uspeh (visoke vrednosti β_{0j}) in šibek učinek DES (nizke vrednosti β_{1j}) – najdemo v zgornjem levem kvadrantu grafikona.

Slika 6.4: PRIKAZ POVPREČJA ŠOL (VERTIKALNO) IN NAGIBOV DRUŽBENO EKONOMSKEGA STATUSA (HORIZONTALNO) ZA 200 HIPOTETIČNIH ŠOL



Vir: Raudenbush in Bryk 2002: 20.

Sedaj, ko smo na ta način grafično pojasnili, kako šole variirajo v smislu presekov in nagibov, velja pojasniti še model, s pomočjo katerega lahko napovedujemo β_{0j} in β_{1j} . V našem specifičnem primeru bi lahko denimo uporabili značilnosti šole (npr. višina investicij oziroma finančnih sredstev, organizacijske značilnosti, politike) za napoved učinkovitosti in enakomernosti porazdelitve. Zato predstavljamo enostaven indikator W_j , ki zavzema vrednost 1 za katoliške šole in vrednost 0 za javne šole. Coleman, Hoffer in Kilgore (v Raudenbush in Bryk 2002: 20) so na ta način trdili, da je W_j pozitivno povezan z učinkovitostjo (katoliške šole imajo višji povprečni uspeh kot javne šole) in negativno povezan z nagibom (učinek DES na matematični uspeh je manjši v katoliških šolah kot pa v javnih). V skladu z zgoraj opisanim lahko ti dve hipotezi predstavimo z dvema regresijskima enačbama:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}W_j + u_{0j} \quad [\text{enačba 6.4a}]$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}W_j + u_{1j}, \quad [\text{enačba 6.4b}]$$

γ_{00} je povprečni uspeh v javnih šolah;

γ_{01} je povprečna razlika v uspehu med katoliškimi in javnimi šolami (npr. prednost v učinkovitosti katoliške šole);

γ_{10} je povprečni nagib odnosa DES-uspešnost v javnih šolah;

γ_{11} je povprečna razlika nagibov odnosa DES-uspešnost med katoliškimi in javnimi šolami (npr. prednost katoliške šole v enakomernosti porazdelitve);

u_{0j} je značilni vpliv šole j na povprečno uspešnost, ki je pogojen s konstanto W_j ;

u_{1j} je značilni vpliv šole j na nagib odnosa DES-uspešnost, ki je pogojen s konstanto W_j .

Predpostavljamo, da sta u_{0j} in u_{1j} naključni spremenljivki s povprečjem 0, variancama τ_{00} in τ_{11} in kovarianco τ_{01} . Te komponente variance-kovariance so zdaj pogojne oziroma rezidualne komponente variance-kovariance. Predstavljajo variabilnost β_{0j} in β_{1j} po tem, ko jih kontroliramo z W_j .

Obenem tudi ni mogoče predvideti parametrov teh regresijskih enačb neposredno, saj izidi (β_{0j} , β_{1j}) niso merjeni. Vendar pa podatki vsebujejo informacije potrebne za predvidevanje le-teh. To postane jasno, če zamenjamo enačbi, ki izhajata iz zgoraj opisanih hipotez, tako da dobimo enotno napovedno enačbo:

$$Y_{ij} = \gamma_{00} + \gamma_{01}W_j + \gamma_{10}(X_{ij} - \mu_{X\cdot j}) + \gamma_{11}W_j(X_{ij} - \mu_{X\cdot j}) + u_{0j} + u_{1j}(X_{ij} - \mu_{X\cdot j}) + r_{ij}$$

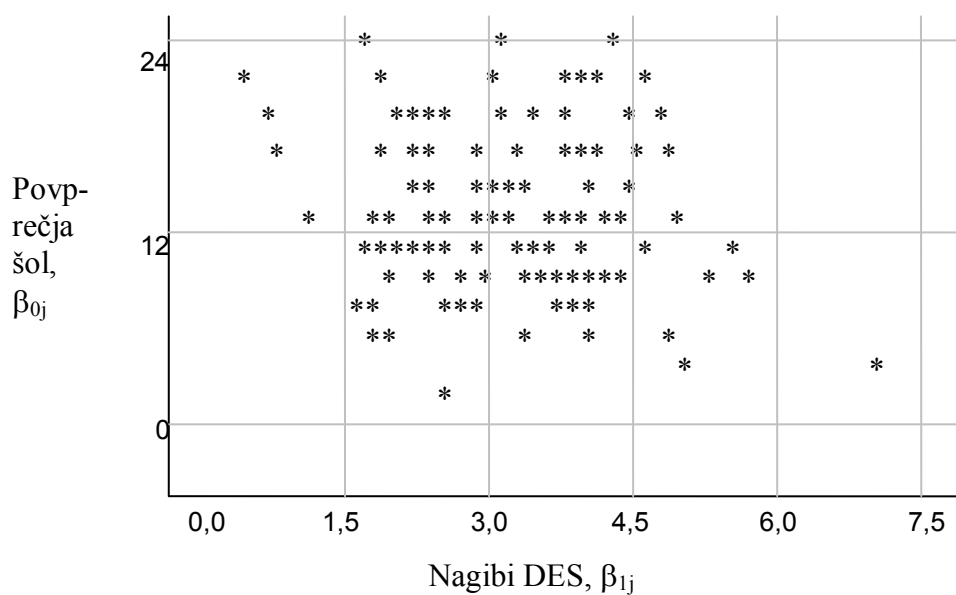
[enačba 6.5]

Kot lahko opazimo, ta enačba ni tipična za linearni model, kot je predpostavljen po vsoti najmanjših kvadratov. Učinkovito in točno sklepanje ter testiranje hipotez, ki temeljita na vsoti najmanjših kvadratov, zahtevata, da so naključne napake neodvisne, normalno porazdeljene, in da imajo konstantno varianco. Po drugi strani pa je naključna napaka, ki jo vsebuje naša enačba, bolj kompleksne oblike: $u_{0j} + u_{1j}(X_{ij} - \mu_{X\cdot j}) + r_{ij}$. Takšne napake so odvisne od vsake šole, saj so komponente u_{0j} in u_{1j} značilne za vsakega študenta v šoli j . Te napake tudi nimajo enakih varianc, ker je rezultat $u_{0j} + u_{1j}(X_{ij} - \mu_{X\cdot j})$ odvisen od u_{0j} in u_{1j} , ki se razlikujeta od šole do šole, in od $(X_{ij} - \mu_{X\cdot j})$, ki se razlikuje od študenta do študenta. Če bi bila u_{0j} in u_{1j} nična za vsak j , potem bi bila zgornja enačba ekvivalentna regresijskemu modelu vsote najmanjših kvadratov. Model, ki smo

ga specificirali z enačbama, ki predstavljata naši hipotezi, je grafično predstavljen na slikah 6.5 in 6.6. Vidimo dva hipotetična grafa odnosa med β_{0j} in β_{1j} , enega za javne šole in drugega za katoliške šole, ki izhajata iz trditev Colemana in drugi, da imajo katoliške šole tako višji povprečni uspeh, kot tudi šibkejši učinek DES kot pa javne šole.

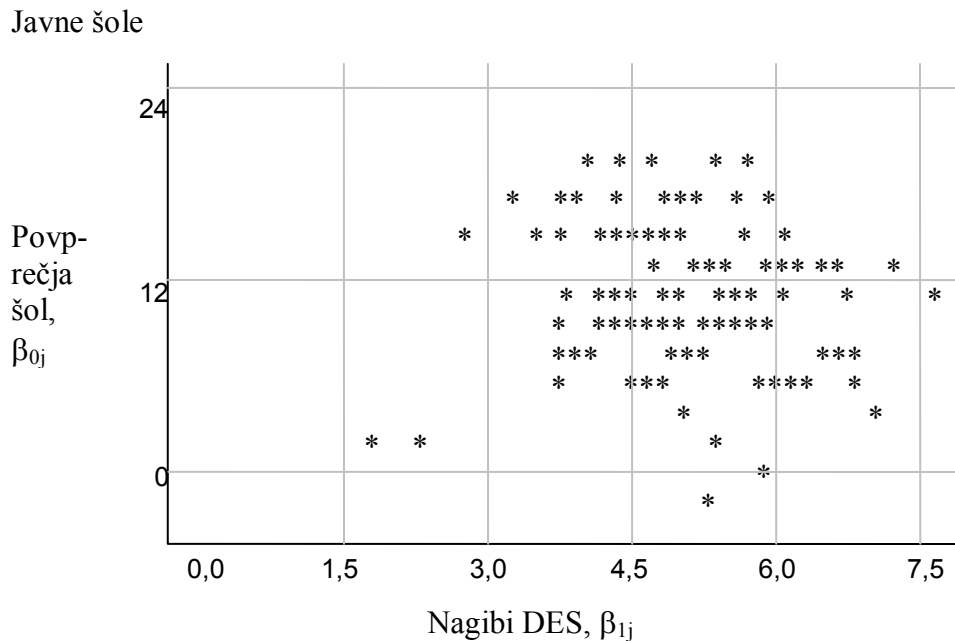
Slika 6.5: GRAF POVPREČIJ ŠOL IN NAGIBOV DRUŽBENO EKONOMSKEGA STATUSA ZA 100 HIPOTETIČNIH KATOLIŠKIH ŠOL

Katoliške šole



Vir: Raudenbush in Bryk 2002: 22.

Slika 6.6: GRAF POVPREČIJ ŠOL IN NAGIBOV DRUŽBENO EKONOMSKEGA STATUSA ZA 100 HIPOTETIČNIH JAVNIH ŠOL



Vir: Raudenbush in Bryk 2002: 22.

Ta princip, ki smo ga opisali skozi poglavje, predstavlja logiko večnivojske hierarhične strukture podatkov. Enačbo $Y_{ij} = \beta_{0j} + \beta_{1j}(X_{ij} - \mu X_{.j}) + r_{ij}$ lahko označimo kot model nivoja 1, enačbi $\beta_{0j} = \gamma_{00} + \gamma_{01}W_j + u_{0j}$ in $\beta_{1j} = \gamma_{10} + \gamma_{11}W_j + u_{1j}$ lahko označimo kot model nivoja 2, enačbo $Y_{ij} = \gamma_{00} + \gamma_{01}W_j + \gamma_{10}(X_{ij} - \mu X_{.j}) + \gamma_{11}W_j(X_{ij} - \mu X_{.j}) + u_{0j} + u_{1j}(X_{ij} - \mu X_{.j}) + r_{ij}$ pa kot kombinirani večnivojski model. Pri apliciranju učinka šole so enote na prvem nivoju študenti, na drugem pa šole. Napake r_{ij} so naključni vplivi nivoja 1, napake u_{0j} in u_{1j} pa naključni učinki nivoja 2. Prav tako je $\text{Var}(r_{ij})$ varianca na prvem nivoju, $\text{Var}(u_{0j})$, $\text{Var}(u_{1j})$ in $\text{Cov}(u_{0j}, u_{1j})$ pa so komponente variance-kovariance na nivoju 2. Parametri β v modelu nivoja 1 so koeficienti prvega nivoja, parametri γ pa so koeficienti drugega nivoja. To pa je obenem tudi logika večnivojskega modeliranja.

Če opisano terminologijo posplošimo še na katerokoli dvonivojsko podatkovno hierarhijo, lahko enačbo 6.2 poimenujemo model nivoja 1, enačbo 6.4 model nivoja 2 ter enačbo 6.5 kot kombiniran model.

Model nivoja 1:

$$Y_{ij} = \beta_{0j} + \beta_{1j}(X_{ij} - \mu_{X.j}) + r_{ij}, \quad [\text{enačba 6.2}]$$

Model nivoja 2:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}W_j + u_{0j} \quad [\text{enačba 6.4a}]$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}W_j + u_{1j}, \quad [\text{enačba 6.4b}]$$

Mešani model:

$$Y_{ij} = \gamma_{00} + \gamma_{01}W_j + \gamma_{10}(X_{ij} - \mu_{X.j}) + \gamma_{11}W_j(X_{ij} - \mu_{X.j}) + u_{0j} + u_{1j}(X_{ij} - \mu_{X.j}) + r_{ij} \quad [\text{enačba 6.5}]$$

(Raudenbush in Bryk 2002)

7. RAZISKOVALNI OKVIR IN ANALIZA

7.1 POJASNJEVALNI MODEL

V namen primerjave večnivojske analize in multiple regresije sem za izhodišče izbrala pojasnjevalni model refleksivnosti članov v spletnih forumih (Petrič 2007). V tem modelu so izpeljane tri vzročno posledične hipoteze, ki opisujejo dejavnike na individualnem nivoju in pet hipotez, ki opisuje dejavnike na nivoju forumov.

H1: Večji kot je občutek pripadnosti posameznika forumu, večja je njegova refleksivnost v razpravah na forumu.

H2: Bolj kot posameznik zaznava norme na forumu, večja je njegova refleksivnost v razpravah na forumu.

H3: Bolj kot je posameznik neanonimen, večja je njegova refleksivnost v razpravah v forumu.

H4: Starejši kot je forum, večja je refleksivnost posameznikov v razpravah na forumu.

H5: Bolj kot je forum izpostavljen, večja je refleksivnost posameznikov v razpravah na forumu.

H6: Če je forum interesen, je refleksivnost posameznikov v razpravah na forumu večja.

H7: Če ima forum jasna pravila, je refleksivnost posameznikov v razpravah na forumu večja.

H8: Če forum zahteva identificiranje članov, je refleksivnost posameznikov v razpravah na forumu večja.

Omenjeni pojasnjevalni model in hipoteze so bile ocenjene z multiplo regresijo fiksnih učinkov (multipla regresijska analiza na deagregiranih podatkih), kar pa ima – kot smo ugotovili v teoretičnem delu – svoje slabosti. Isti model sem torej preverila še z večnivojsko analizo in rezultate primerjala z regresijsko analizo.

7.2 OPIS VZORCA IN ZBIRANJA PODATKOV

Pojasnjevalni model smo preverjali na istih podatkih kot ko je bil model preverjan z regresijsko analizo (Petrič 2007). Podatki za raziskavo so bili zbrani s pomočjo spletne ankete, ki je bila del raziskovalnega projekta "Družbeni potencial slovenskih spletnih forumov", ki se je izvajal na Centru za metodologijo in informatiko na FDV. Spletna anketa, ki je gostovala na spletnem naslovu podjetja Cati d.o.o. ankete.cati.si/fdv-splet2, se je oglaševala preko različnih spletnih strani, predvsem na forumih. V raziskavo je bilo vključenih 1895 uporabnikov različnih forumov.

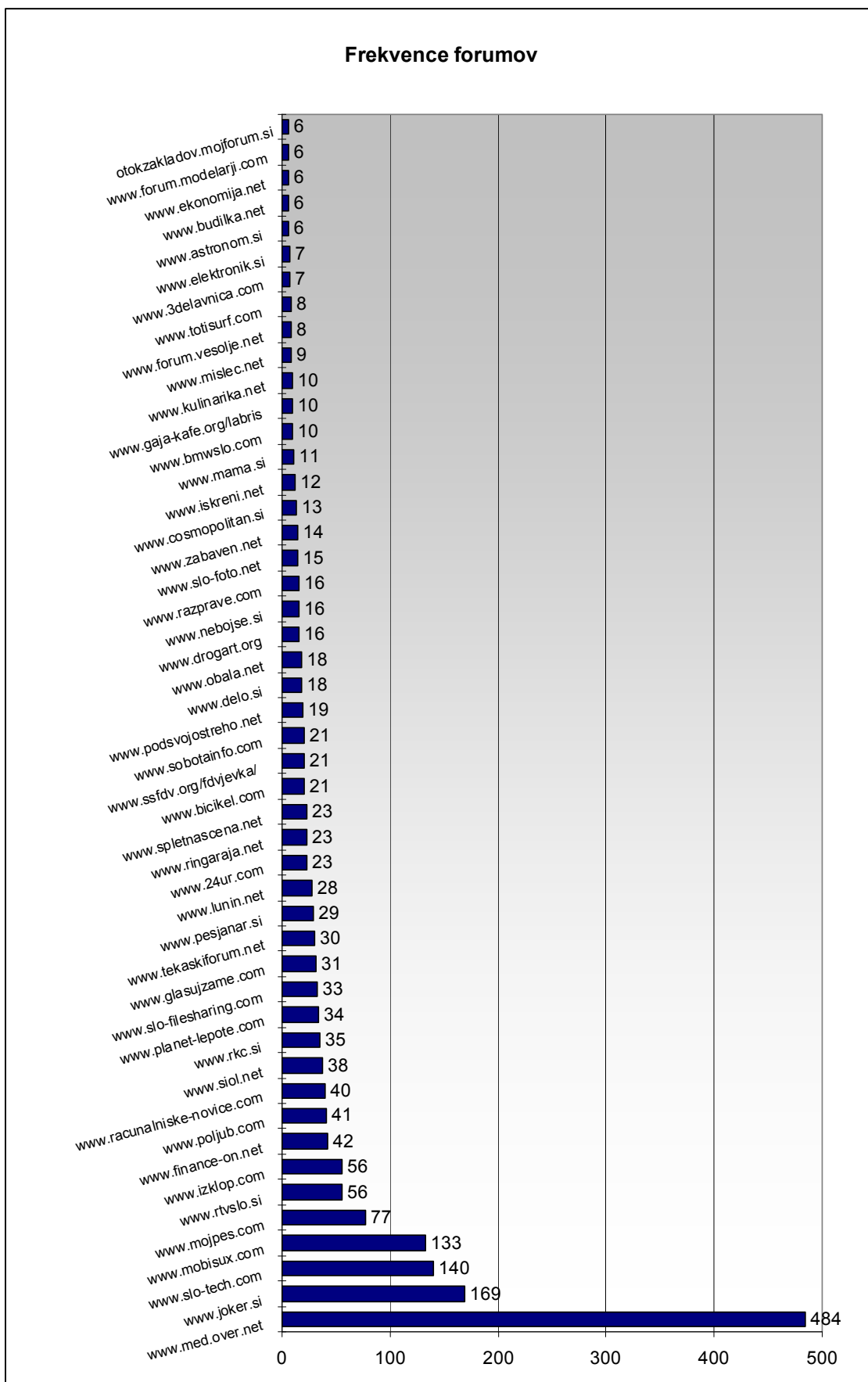
Tabela 7.1: OPIS VZORCA IN STRUKTURA UPORABNIKOV –
SOCIODEMOGRAFSKI PODATKI

Spremenljivka	Vrednost	n	%
SPOL	Moški	924	48.9
	Ženski	966	51.1
STAROST	10-19 let	195	10.7
	20-29 let	837	45.8
	30-49 let	723	39.6
	50-65 let	67	3.7
	65+ let	5	0.3
IZOBRAZBA	Nepopolna osnovna izobrazba	31	1.6
	Osnovna izobrazba	202	10.7
	Nižja ali srednja poklicna izobrazba	43	2.3
	Srednja strokovna izobrazba	305	16.2
	Srednja splošna izobrazba	605	32.2
	Višja strokovna izobrazba, višješolska izobrazba	136	7.2
	Visoka strokovna izobrazba	147	7.8
	Visoka univerzitetna izobrazba	338	18.0
	Specialistična povisokošolska izobrazba, magisterij, doktorat	74	3.9
Kako pogosto ste v povprečju uporabljali internet v zadnjih 3 mesecih?	Vsak dan ali skoraj vsak dan	1830	96.7
	Vsaj enkrat na teden (vendar ne vsak dan)	55	2.9
	Vsaj enkrat na mesec (vendar ne vsak teden)	5	0.3
	Manj kot enkrat na mesec	2	0.1
	Ne vem	1	0.1

Glede na demografske podatke lahko ugotovimo, da je v raziskavi sodelovalo 49 odstotkov moških in 51 odstotkov žensk. Glede na porazdelitev

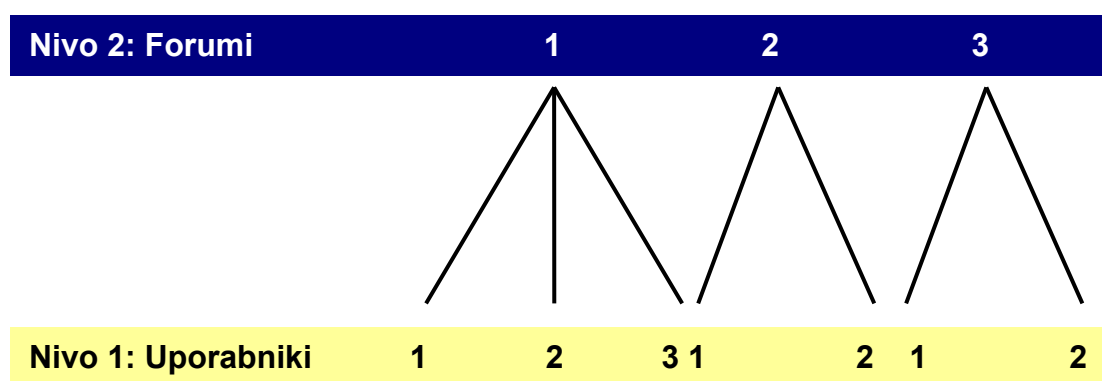
respondentov na starostne razrede ugotovimo, da je v raziskavi sodelovalo največ anketirancev starih med 20 in 29 let (46%), nekoliko manj starih med 30 in 49 let (40%), sledijo jim respondenti stari od 10 do 19 let z 11 odstotki ter 50 do 65 let s 4 odstotki. Pri razdelitvi respondentov glede na doseženo stopnjo izobrazbe ugotovimo, da je največ respondentov s srednjo splošno izobrazbo (32%), nekoliko manj je takšnih z visoko univerzitetno izobrazbo (18%) in s srednjo strokovno izobrazbo (16%). Najmanj respondentov navaja, da nima izobrazbe oziroma imajo nepopolno osnovno izobrazbo (2%). Anketirance smo povprašali tudi, kako pogosto so uporabljali internet v zadnjih treh mesecih. Z veliko večino prevladujejo respondenti, ki so v zadnjih treh mesecih uporabljali internet vsak dan ali skoraj vsak dan. Takšnih respondentov je kar 97 odstotkov.

Slika 7.1: FREKVENCE FORUMOV

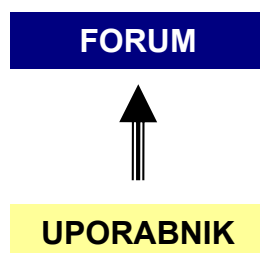


Ko pregledamo uporabnike glede na forum, v katerem sodelujejo, opazimo različno število uporabnikov v vsakem forumu. Ko te enote razporedimo glede na večnivojsko analizo, lahko opazimo, da imamo 48 enot nivoja 2 ter 1895 enot nivoja 1. Forumi namreč predstavljajo enote nivoja 2, v katerih so združeni uporabniki, ki predstavljajo enote nivoja 1. Za lažjo preglednost so spremenljivke nižjega nivoja⁹ obarvane z rumeno in spremenljivke višjega nivoja¹⁰ z modro barvo.

Slika 7.2: DIAGRAM ENOT



Slika 7.3: KLASIFIKACIJSKI DIAGRAM



⁹ Pojem se nanaša na spremenljivke, ki karakterizirajo posameznike oziroma enote nižjega nivoja ter na konstrukte nižjega nivoja.

¹⁰ Pojem se uporablja za spremenljivke, ki karakterizirajo skupine. Podobni pojmi so makro spremenljivke ali ekološke spremenljivke (Roux 2006). Takšne spremenljivke lahko uporabljamo kot neke vrste nadomestnike za nedosegljive ali nezanesljive podatke nižjega nivoja, v večnivojski analizi pa predvsem kot kazalce sestavin višjega nivoja.

7.3 INDIKATORJI

V teoretičnem pojasnjevalnem modelu je 9 teoretičnih spremenljivk (ena odvisna in 8 neodvisnih), na empiričnem nivoju pa se pokažejo za določene neodvisne spremenljivke poddimenzije, zato je v končni fazi v modelu 9 neodvisnih spremenljivk na individualnem nivoju in 5 neodvisnih spremenljivk na forumskem nivoju (Petrič 2007).

Tabela 7.2: OPISI ODVISNE IN NEODVISNIH SPREMENLJIVK UPORABLJENIH V RAZISKAVI

Ime spremenljivke	Opis
<i>Odvisna spremenljivka</i>	
Refleksivnost	q13e Skozi pogovor dobim jasnejšo sliko o temi, o kateri razpravljamo q13h Razmišljam o mnenjih drugih udeležencev foruma
<i>Neodvisne spremenljivke nivoja 1</i>	
Občutek pripadnosti	q18a Člani foruma si delimo podobna prepričanja in vrednote. q18c V forumu se počutim zelo domače. q18e Sodelovanje v tem forumu mi pomeni zelo veliko v mojem življenju. q18f Člani foruma se med seboj dobro poznamo q18g Če pride do problemov v forumu, se vsi trudimo, da jih odpravimo q18i Člani foruma se dobro razumemo
Percepcija solidarnosti	q19a Na forumu se objavljajo koristne informacije q19b V forumu si sodelujoči pomagajo z nasveti q19f Na forumu se ljudje spoznavajo in družijo
Prežemanje z offline	q15 Približno koliko oseb od vseh sodelujočih v forumu, poznate tudi osebno (v offline življenju)?
Formalna pravila	q20 Ali ima forum javno objavljena pravila vedenja in delovanja?
Formalne sankcije	q21a V forumu imajo določeni člani pravico, da popravijo ali izbrišejo sporočila drugih q21b Člani foruma opozarjajo druge udeležence na neprimerno vedenje in izražanje q21d V forumu so moderatorji, ki kaznujejo ali izločajo uporabnike, ki se ne držijo pravil q21f Žaljiva in neprimerna sporočila moderatorji izbrišejo ali popravijo
Neformalne sankcije	q21e V forumu se je potrebno držati določenih nenapisanih pravil q21g Članom foruma je jasno, kakšen način komuniciranja je zaželen q21k Člani foruma ignorirajo sporočila, ki odstopajo od »forumskega bontona«.

Pozitivne sankcije	<p>S primernim vedenjem se da v forumu doseči status pomembnega člana.</p> <p>Člani foruma spodbujajo k pisanju sporočil, ki so koristna za cel forum.</p> <p>Če v forumu nekdo piše preudarno, ga ostali člani spoštujejo</p>
Identifikabilnost	<p>Ali vi v forumu, v katerem ste najbolj aktivni, uporabljate/podajate</p> <p>q21h avatar s svojo osebno fotografijo</p> <p>q21i podpis, ki vključuje vaš e-mail, ali povezavo na osebno predstavitevno mesto (blog)</p> <p>q21j resnične osebne podatke v profilu (npr. spol, starost, kraj bivanja, MSN naslov)</p> <p>q10e svoje resnično osebno ime</p> <p>q10g vzdevek (nick-name), ki ga uporabljam tudi v vsakdanjem življenju</p> <p>q10f podpis z razpoznavnim grafičnim ali besedilnim motivom</p> <p>q10a vzdevek (nick-name), ki ne spominja na moje ime, priimek ali prostor, kjer živim</p>
Koherentnost samopredstavitve	<p>Ali ste, odkar sodelujete v forumu, spremenili, kaj od naslednjega</p> <p>q11a svoj vzdevek</p> <p>q11b svoj avatar</p> <p>q11c podatke v osebni profilu</p> <p>q11d svoj podpis z razpoznavnim grafičnim ali besedilnim motivom</p> <p>q11e podpis, ki vključuje vaš e-mail, ali povezavo na osebno predstavitevno stran (blog)</p>

Neodvisne spremenljivke nivoja 2

Starost	Starost foruma
Rank Alexa	Rang popularnosti, ki ga zaseda spletno mesto, kjer se nahaja forum, med vsemi slovenskimi spletnimi mesti
Tip1	Interesnost foruma (ustanovljen s strani članov ali organizacije)
Pravila	Prisotnost pravil v forumu
Identifikacija	Stopnja obvezne identifikacije, ki jo zahteva forum

7.4 PRIMERJAVA OPISNIH STATISTIK

V tem podglavju želim predstaviti in primerjati opisne statistike pravih večnivojskih podatkov in opisne statistike spremenljivk, ki nastopajo v regresiji s fiksnimi učinki.

Tabela 7.3: OPISNE STATISTIKE V REGRESIJI

Spremenljivka	n	Povprečje	SD	Minimum	Maksimum
Starost	1895	6.16	2.26	2.00	10.00
Rank Alexa	1895	336.08	777.08	5.00	8579.00
Tip1	1895	1.55	0.50	1.00	2.00
Pravila	1895	2.33	0.89	0.00	4.00
Identifikacija	1895	0.44	0.75	0.00	2.00
Refleksivnost	1849	3.70	0.73	1.00	5.00
Občutek pripadnosti	1824	2.94	0.72	1.00	5.00
Percepcija solidarnosti	1821	3.96	0.61	1.00	5.00
Prežemanje z offline	1885	2.07	1.75	1.00	9.00
Formalna pravila	1852	3.43	0.77	1.00	4.00
Formalne sankcije	1818	4.09	1.28	1.00	9.00
Neformalne sankcije	1811	3.40	1.30	1.00	9.00
Pozitivne sankcije	1809	4.23	1.39	1.00	9.00
Identifikabilnost	1796	3.29	2.69	0.00	14.00
Koherentnost samopredstavitve	1815	8.29	1.91	0.00	10.00

Tabela 7.4: OPISNE STATISTIKE HIERARHIČNIH PODATKOV ZA NIVO 2

OPISNE STATISTIKE NIVOJA 2					
Spremenljivka	n	Povprečje	SD	Minimum	Maksimum
Starost	48	4.90	1.82	2.00	10.00
Rank Alexa	48	856.98	1446.82	5.00	8579.00
Tip1	48	1.48	0.50	1.00	2.00
Pravila	48	2.19	0.94	0.00	4.00
Identifikacija	48	0.50	0.74	0.00	2.00

Tabela 7.5: OPISNE STATISTIKE HIERARHIČNIH PODATKOV ZA NIVO 1

OPISNE STATISTIKE NIVOJA 1						
Spremenljivka	n	Povprečje		SD	Minimum	Maksimum
Refleksivnost	1849	3.70	0.73		1.00	5.00
Občutek pripadnosti	1824	2.94	0.72		1.00	5.00
Percepcija solidarnosti	1821	3.96	0.61		1.00	5.00
Prežemanje z offline	1885	2.07	1.75		1.00	9.00
Formalna pravila	1852	3.43	0.77		1.00	4.00
Formalne sankcije	1818	4.09	1.28		1.00	9.00
Neformalne sankcije	1811	3.40	1.30		1.00	9.00
Pozitivne sankcije	1809	4.23	1.39		1.00	9.00
Identifikabilnost	1796	3.29	2.69		0.00	14.00
Koherentnost samopredstavitve	1815	8.29	1.91		0.00	10.00

Iz tabel 7.3, 7.4 in 7.5 opazimo, da med spremenljivkami preučevanimi z dvema različnima metodama prihaja do velikih razhajanj. Do teh razhajanj pride zaradi pristranskosti regresije, ki ne upošteva hierarhične strukture podatkov, kar se odraža predvsem v spremenljivkah, ki merijo različne lastnosti enot nivoja 2, torej forumov. V prvi vrsti gre izpostaviti število enot, n, pri spremenljivkah drugega nivoja. V regresijski analizi smo za enote preučevanje uporabili uporabnike forumov, zaradi tega je število preučevanih enot, pri spremenljivkah drugega nivoja 1895. Po drugi strani je to število enot v večnivojski analizi občutno manjše, saj enote analize predstavljajo forumi, katerih je 48. Zaradi različnega števila enot, ki jih upoštevata obe analizi prihaja do večjih razlik tudi pri drugih opisnih statistikah teh spremenljivk, in sicer pri povprečju in standardnem odklonu. Povprečna starost forumov je glede na regresijsko analizo 6,16 let, med tem ko je le ta v primeru večnivojske analize 4,90 let. Takšna razlika v povprečni starosti foruma je ravno posledica različnega števila enot. Do podobnih razlik prihaja tudi v primeru ostalih štirih pojasnjevalnih spremenljivk na drugem nivoju. Povprečje spremenljivke Rank Alexa oziroma popularnosti foruma je v primeru regresije 336,08, v primeru upoštevanja hierarhične strukture podatkov pa 856,89. Povprečje spremenljivke Tip1 oziroma interesnost foruma je v prvem primeru 1,55 in v drugem 1,48, povprečje spremenljivke pravila je v prvem primeru 2,33 in v drugem 2,19. ter nazadnje še povprečje spremenljivke identifikacija, ki je v prvem primeru 0,44 v drugem pa 0,50.

Opisne spremenljivke za prvi nivo, nivo uporabnikov, so enake v obeh primerih. Obe metodi namreč v tem primeru upoštevata enako število enot, zaradi česar so povprečja na vseh opisnih spremenljivkah nivoja ena enake.

7.5 REZULTATI REGRESIJE S FIKSNIMI UČINKI

Kot sem že omenila sem regresijsko analizo izvedla na deagregiranih podatkih, torej sem za enote analize uporabila uporabnike forumov. Enot analize je skupno 1895, ki sodelujejo v 48 različnih forumih. Regresijsko analizo sem izvedla v statističnem paketu SPSS in dobila naslednje rezultate.

Determinacijski koeficient R znaša 0,330 in nam pove, koliko odstotkov variabilnosti odvisne spremenljivke lahko pojasnimo z neodvisno spremenljivko. Glede na vrednost R^2 lahko ugotovimo, da model pojasnjuje 10,9 odstotkov variabilnosti. Pojem napake v multipli regresiji meri razliko med posameznikovo reflektivnostjo in povprečno reflektivnostjo posameznika z istimi lastnostmi. F-test za nično hipotezo, po kateri nobena od pojasnjevalnih spremenljivk ni povezana z reflektivnostjo, nam pokaže, da lahko nično hipotezo z gotovostjo zavrnemo saj je $F=14,346$ in $p < 0,001$ in na podlagi tega lahko sklepamo, da je vsaj ena od neodvisnih spremenljivk povezana z reflektivnostjo.

Tabela 7.6: REGRESIJSKI KOEFICIENTI^a

	Nestandardizirani koeficienti		Standardizirani koeficienti		Sig.
	Beta	Std.napaka	Beta	t	
Konstanta	0.005	0.023		0.228	0.820
Starost	-0.020	0.035	-0.020	-0.565	0.572
Rank Alexa	0.045	0.025	0.047	1.771	0.077
Tip1	-0.006	0.034	-0.006	-0.187	0.852
Pravila	0.052	0.024	0.053	2.165	0.031
Identifikacija	0.029	0.024	0.029	1.225	0.221
Občutek pripadnosti	0.159	0.029	0.161	5.532	0.000
Percepcija solidarnosti	0.158	0.028	0.156	5.611	0.000
Prežemanje z offline	0.013	0.023	0.013	0.574	0.566
Formalna pravila	0.092	0.026	0.093	3.581	0.000
Formalne sankcije	0.018	0.027	0.019	0.679	0.497
Neformalne sankcije	-0.066	0.029	-0.068	-2.282	0.023
Pozitivne sankcije	0.016	0.028	0.017	0.597	0.551
Identifikabilnost	-0.013	0.025	-0.013	-0.524	0.601
Koherentnost samopredstavitve	0.009	0.024	0.009	0.364	0.716

a. Odvisna spremenljivka: Refleksivnost

Tabela 7.7: POVZETEK MODELA

Model	Prilagojen R^2	Sig
	0.101	0.000

Iz tabele 7.6 je razvidno, da je med strukturiranimi spremenljivkami oziroma spremenljivkami, ki označujejo lastnosti forumov, le ena spremenljivka statistično povezana z odvisno spremenljivko reflektivnost. To je spremenljivka pravila, ki pozitivno korelira z odvisno spremenljivko, kar pomeni, da bolj kot ima forum zastavljena jasna pravila, večja je reflektivnost posameznikov v razpravah na forumu. S tem lahko potrdimo hipotezo H7. Po drugi strani pa je nekaj več spremenljivk na individualnem nivoju, ki korelirajo z odvisno spremenljivko. Takšne neodvisne spremenljivke, ki izražajo lastnosti posameznikov na forumih so štiri in sicer občutek pripadnosti, percepcija solidarnosti, formalna pravila ter neformalne sankcije. Glede na predznake koeficientov ugotovimo, da občutek pripadnosti forumu pozitivno korelira z reflektivnostjo. Povedano z drugimi besedami, večji kot je občutek pripadnosti posameznika forumu, večja je njegova reflektivnost v razpravah na forumu. S tem potrdimo hipotezo H1. Tudi percepcija solidarnosti pozitivno korelira z reflektivnostjo. Torej, večja kot je percepcija solidarnosti, večja je reflektivnost v razpravah v forumu. Z reflektivnostjo pozitivno korelira tudi neodvisna spremenljivka zaznavanje formalnih pravil. S tem lahko še na nivoju posameznikov potrdimo hipotezo H7: če ima forum jasna pravila, je reflektivnost posameznikov v razpravah na forumu večja. Nekoliko drugače pa se izkaže za neodvisno spremenljivko neformalne sankcije, ki je, glede na predznak koeficienta, negativno povezana z reflektivnostjo. Večja kot je torej zaznava neformalnih sankcij s strani uporabnikov forumov, manjša je reflektivnost v razpravah na forumu.

7.6 REZULTATI VEČNIVOJSKE ANALIZE

Večnivojsko analizo sem izvedla s pomočjo statističnega orodja HLM. Najprej sem natančno opredelila neodvisne spremenljivke glede na nivo, ki mu pripadajo. Enote analize na višjem nivoju predstavljajo forumi, ki jih je 48, enote analize na nižjem nivoju pa uporabniki forumov, ki jih je 1895. Takšna opredelitev spremenljivk nam poda model nivoja ena ter model nivoja dva razložena spodaj.

Model nivoja-1 oziroma nivoja uporabnikov

Pri opredelitvi modela nivoja 1 sem označila odvisno spremenljivko za uporabnika i v forumu j kot Y_{ij} . Ta odvisna spremenljivka je predstavljena kot funkcija lastnosti uporabnikov, X_{qij} in napake modela r_{ij} :

$$Y_{ij} = \beta_{01} + \beta_1 X_{1ij} + \beta_2 X_{2ij} + \beta_3 X_{3ij} + \beta_4 X_{4ij} + \beta_5 X_{5ij} + \beta_6 X_{6ij} + \beta_7 X_{7ij} + \beta_8 X_{8ij} + \beta_9 X_{9ij} + r_{ij}$$

$$\begin{aligned} \text{Refleksivnost}_{ij} = & \beta_{01} + \beta_{1j}^*(\text{ObčutekPripadnosti})_{ij} + \beta_{2j}^*(\text{PercepcijaSolidarnosti})_{ij} \\ & + \beta_{3j}^*(\text{PrežemanjeZoffline})_{ij} + \beta_{4j}^*(\text{FormalnaPravila})_{ij} + \beta_{5j}^*(\text{FormalneSankcije})_{ij} \\ & + \beta_{6j}^*(\text{NeformalneSankcije})_{ij} + \beta_{7j}^*(\text{PozitivneSankcije})_{ij} + \beta_{8j}^*(\text{Identifikabilnost})_{ij} \\ & + \beta_{9j}^*(\text{KoherentnostSamopredstavitve})_{ij} + r_{ij} \end{aligned}$$

β_{01} = povprečna vrednost refleksivnosti v forumu j

V tem primeru uvodoma predpostavljamo, da $r_{ij} \sim N(0, \delta^2)$

Regresijski koeficienti β_{q1} , $q = 0, \dots, Q$ pokažejo, kako je rezultat razporejen v forumu j kot funkcija merjenih lastnosti uporabnikov. Imenujemo jih distributivni učinki (Raudenbush, Bryk 2002).

Model nivoja-2 oziroma nivoja forumov

Za učinke vsakega foruma, ki so zajeti v nabor β_{qj} v prejšnji enačbi, predpostavljamo, da variirajo po enotah. Vsak β_{qj} ima model naslednje oblike:

$$\beta_{qj} = \gamma_{q0} + \gamma_{q1} W_{1j} + \gamma_{q2} W_{2j} + \gamma_{q3} W_{3j} + \gamma_{q4} W_{4j} + \gamma_{q5} W_{5j} + u_{qj}$$

$$\beta_{1j} = \gamma_{10} + u_{1j}$$

$$\beta_{2j} = \gamma_{20} + u_{2j}$$

$$\beta_{3j} = \gamma_{30} + u_{3j}$$

$$\beta_{4j} = \gamma_{40} + u_{4j}$$

$$\beta_{5j} = \gamma_{50} + u_{5j}$$

$$\beta_{6j} = \gamma_{60} + u_{6j}$$

$$\beta_{7j} = \gamma_{70} + u_{7j}$$

$$\beta_{8j} = \gamma_{80} + u_{8j}$$

$$\beta_{9j} = \gamma_{90} + u_{9j}$$

$$B_{0j} = \gamma_{00} + \gamma_{01}*(Starost) + \gamma_{02}*(RankAlexa) + \gamma_{03}*(Tip1) + \gamma_{04}*(Pravila) + \gamma_{05}*(Identifikacija) + u_{0j}$$

Mešani model

$$\begin{aligned} Refleksivnost = & \gamma_{00} + \gamma_{01}*(Starost) + \gamma_{02}*(RankAlexa) + \gamma_{03}*(Tip1) + \gamma_{04}*(Pravila) \\ & + \gamma_{05}*(Identifikacija) + \gamma_{10}*(ObčutekPripadnosti)_{ij} + \gamma_{20}*(PercepcijaSolidarnosti)_{ij} \\ & + \gamma_{30}*(PrežemanjeZoffline)_{ij} + \gamma_{40}*(FormalnaPravila)_{ij} + \gamma_{50}*(FormalneSankcije)_{ij} \\ & + \gamma_{60}*(NeformalneSankcije)_{ij} + \gamma_{70}*(PozitivneSankcije)_{ij} + \gamma_{80}*(Identifikabilnost)_{ij} \\ & + \gamma_{90}*(KoherentnostSamopredstavitve)_{ij} + u_{0j} + r_{ij} \end{aligned}$$

Tabela 7.8: OCENE DELEŽEV NAKLJUČNIH UČINKOV

Sigma2	Tau Presek1, β_0	Presek1, β_0
0.45354	0.00892	1.000

Tabela prikazuje ocene deleža naključnih učinkov v modelu. Z uporabo večnivojskega modela lahko celotno varianco razdelimo na varianco znotraj skupin in varianco med skupinami. Ocena variance nivoja 1 oziroma znotraj forumov znaša $\text{Var}(r_{ij}) = \delta^2 = 0,45354$, ki nam pove kakšna je variabilnost med uporabniki forumov. Ocena variance nivoja 2 oziroma populacijske variance za aritmetične sredine forumov znaša $\text{Var}(u_{0j}) = \tau_{00} = 0,00892$, ki nam pove, kolikšna je variabilnost med forumi. Na podlagi teh dveh podatkov lahko izračunamo znotrajrazredno korelacijo¹¹, ki znaša 0,019288154. Ta podatek

¹¹ $p = \tau_{00} / (\tau_{00} + \delta^2) = 0,00892 / (0,00892 + 0,45354) = 0,019288154$

nam pove, da 1,9 odstotka variance refleksivnosti lahko pojasnimo z razlikami med forumi.

Tabela 7.9: KOEFICIENTI

Fiksni učinki	Koeficienti	Standardne napake	t	n	p-vrednost
Za Presek1, β_0 in presek2 γ_{00}	3.791068	0.084712	44.753	42	0.000
Starost γ_{01}	-0.008403	0.012419	-0.677	42	0.502
Rank Alexa γ_{02}	0.000046	0.00002	2.254	42	0.029
Tip1 γ_{03}	-0.046113	0.054033	-0.853	42	0.398
Pravila γ_{04}	0.061398	0.028253	2.173	42	0.035
Identifikacija γ_{05}	0.028679	0.026118	1.098	42	0.279
Za Občutek pripadnosti nagib, β_1 in presek γ_{10}	0.162302	0.032034	5.067	1643	0.000
Za Percepcija solidarnosti nagib, β_2 in presek γ_{20}	0.199766	0.039625	5.041	1643	0.000
Za Prežemanje z offline nagib, β_3 in presek γ_{30}	0.006883	0.010173	0.677	1643	0.499
Za Formalna pravila nagib, β_4 in presek γ_{40}	0.09832	0.020708	4.748	1643	0.000
Za Formalne sankcije nagib, β_5 in presek γ_{50}	0.011765	0.016204	0.726	1643	0.468
Za Neformalne sankcije nagib, β_6 in presek γ_{60}	-0.037623	0.019441	-1.935	1643	0.053
Za Pozitivne sankcije nagib, β_7 in presek γ_{70}	0.008389	0.013742	0.610	1643	0.541
Za Identifikabilnost nagib, β_8 in presek γ_{80}	-0.001053	0.008048	-0.131	1643	0.896
Za Koherentnost samopredstavitve nagib, β_9 in presek γ_{90}	0.000548	0.009352	0.059	1643	0.954

Iz tabele je razvidno, da nekatere neodvisne spremenljivke statistično pomembno vplivajo na odvisno spremenljivko. Ugotovimo lahko, da na refleksivnost vplivajo rank alexa, pravila, občutek pripadnosti, percepcija solidarnosti in formalna pravila. Med neodvisnimi spremenljivkami višjega nivoja se za statistično značilni izkažeta dve spremenljivki, rank alexa ter pravila. Potrdimo lahko torej hipotezi H5: bolj kot je forum izpostavljen, večja je refleksivnost posameznikov v razpravah na forumu in H7: če ima forum jasna pravila, je refleksivnost posameznikov v razpravah na forumu večja. Med

neodvisnimi spremenljivkami nižjega nivoja, ki statistično pomembno vplivajo na odvisno spremenljivko reflektivnost, so občutek pripadnosti, percepcija solidarnosti in formalna pravila. S tem lahko potrdimo hipotezo H1: večji kot je občutek pripadnosti posameznika forumu, večja je njegova reflektivnost v razpravah na forumu. Trdimo lahko tudi, da večja kot je uporabnikova percepcija solidarnosti, večja je reflektivnost v razpravah na forumu ter večja kot je zaznava formalnih pravil, večja je reflektivnost v razpravah na forumu.

7.7 PRIMERJAVA DVEH ANALITIČNIH PRISTOPOV

Na tem mestu prikazujem primerjavo rezultatov regresije s fiksnimi učinki in večnivojske analize. Osnovni koncept večnivojskega modeliranja je podoben regresiji najmanjših kvadratov, predvsem na nivoju 1, kjer odvisno spremenljivko predstavlja funkcija linearne kombinacije ene ali večih spremenljivk nivoja 1 ter preseka:

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_1 + \dots + \beta_{kj}X_k + r_{ij},$$

kjer β_{0j} predstavlja povprečje za skupino j , β_{1j} predstavlja nagib za spremenljivko X_1 v skupini j in r_{ij} predstavlja rezidual za posameznika i znotraj skupine j . Pri naknadnih nivojih postanejo nagibi in preseki nivoja 1 odvisni od spremenljivk nivoja 2:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}W_1 + \dots + \gamma_{0k}W_k + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}W_1 + \dots + \gamma_{1k}W_k + u_{1j}, \dots,$$

kjer sta γ_{00} in γ_{10} preseka, γ_{01} in γ_{11} nagiba, ki napovedujeta β_{0j} , β_{1j} ločeno od spremenljivke W_1 . S takšnim procesom natančno modeliramo učinke spremenljivk nivoja 1 in učinke spremenljivk nivoja 2 na odvisno spremenljivko. Po tem, ko lahko napovedujemo nagibe in preseke oziroma povprečja, lahko nadalje modeliramo čeznivojske interakcije s čimer lahko poskušamo razumeti, kaj pojasnjuje razlike v povezavah med spremenljivkami nivoja 1 ter rezultati (Osborne 2000).

Za prikaz rezultatov in primerjavo dveh različnih analitičnih pristopov k podatkom s hierarhično strukturo nam bo služila tabela 7.10, ki predstavlja povzetek analiz predstavljenih v prejšnjih dveh podpoglavjih.

Tabela 7.10: PRIMERJAVA DVEH STATISTIČNI PRISTOPOV¹²

Spremenljivka	DEAGREGACIJA			VEČNIVOJSKA ANALIZA		
	B	SE	t	B	SE	T
Starost	-0.020	0.035	-0.565	-0.008	0.012	-0.677
Rank alexa	0.045 *	0.025	1.771	0.000 **	0.000	2.254
Tip1	-0.006	0.034	-0.187	-0.046	0.054	-0.853
Pravila	0.052 **	0.024	2.165	0.061 **	0.028	2.173
Identifikacija	0.029	0.024	1.225	0.029	0.026	1.098
Občutek pripadnosti	0.159 ***	0.029	5.532	0.162 ***	0.032	5.067
Percepcija solidarnosti	0.158 ***	0.028	5.611	0.200 ***	0.040	5.041
Prežemanje z offline	0.013	0.023	0.574	0.007	0.010	0.677
Formalna pravila	0.092 ***	0.026	3.581	0.098 ***	0.021	4.748
Formalne sankcije	0.018	0.027	0.679	0.012	0.016	0.726
Neformalne sankcije	-0.066 **	0.029	-2.282	-0.038 *	0.019	-1.935
Pozitivne sankcije	0.016	0.028	0.597	0.008	0.014	0.610
Identifikabilnost	-0.013	0.025	-0.524	-0.001	0.008	-0.131
Koherentnost samopredstavitve	0.009	0.024	0.009	0.001	0.009	0.059

Na tem mestu prikazujem primerjavo rezultatov regresije s fiksnimi učinki in večnivojske analize. Za izvršitev regresije s fiksnimi učinki sem izvedla deagregacijo podatkov, kjer sem vse vrednosti spremenljivk nivoja 2 pripisala enotam nivoja 1 torej uporabnikom forumov glede na to, kateri enoti nivoja 2 pripadajo. Regresijo sem izvedela v statističnem paketu SPSS, kjer sem sočasno vnesla vse neodvisne spremenljivke. Kot že rečeno, se je model regresijske analize pokazal kot signifikanten z naslednjimi parametri:

$R = 0,330$, $R\text{-square} = 0,109$, $F = 14,346$, $p < ,000$.

Z neodvisnimi spremenljivkami lahko pojasnimo 10,9 odstotkov variabilnosti. Izkazalo se je, da pet neodvisnih spremenljivk statistično značilno vpliva na odvisno spremenljivko reflektivnost. Od teh petih spremenljivk so štiri takšne, ki pozitivno vplivajo na uporabnikovo reflektivnost v razpravah na forumih. To so prisotnost pravil v forumu, uporabnikov občutek pripadnosti, percepcija solidarnosti ter zaznavanje formalnih pravil. Na eni od neodvisnih spremenljivk pa ugotovimo negativen učinek na reflektivnost in sicer v primeru zaznave neformalnih sankcij. Izkaže se torej, da se z večjim zaznavanjem neformalnih sankcij zmanjšuje stopnja reflektivnosti uporabnikov pri v razpravah na forumu.

¹² B se nanaša na nestandardizirane regresijske koeficiente in se uporabljajo v večnivojski analizi za predstavitev nestandardiziranih regresijskih koeficientov, kljub temu da so ponavadi označene z betami in gamami. SE se nanaša na standardno napako (*angl. standard error*). Zvezdice zraven B-jev predstavljajo različno moč signifikanc glede na: * $0,1 > \text{sig} \leq 0,05$; ** $0,05 > \text{sig} \geq 0,01$; *** $\text{sig} < 0,01$.

Večnivojsko analizo sem izvedla v statističnem paketu HLM, kjer sem natančno opredelila odvisne spremenljivke glede na nivo, ki mu pripadajo. Mera prilagojenosti (angl: goodness of fit) modela podatkom je Hi-kvadrat in znaša 64,74 pri čemer je njegova statistična značilnost $p = 0,014$. V HLM modelu ni na voljo statistike, ki bi izražala R ali R-square, lahko pa rečemo, da gre v obeh primerih za modela, ki se statistično značilno prilegata podatkom. Analiza razkriva statistično značilen pozitiven vpliv nekaterih neodvisnih spremenljivk nivoja 1 na reflektivnost – občutek pripadnosti, percepcija solidarnosti in formalna pravila. Opazimo lahko tudi signifikantno pozitiven vpliv spremenljivk nivoja 2 na reflektivnost in sicer spremenljivk Rank Alexa oziroma popularnost foruma in pravila oziroma prisotnost pravil v forumu. Glede na izračun znotrajrazredne korelacije, ki znaša 0,019288154, lahko ugotovimo, da 1,9 odstotka variance v reflektivnosti lahko pojasnimo z razlikami med forumi.

Opazimo torej lahko, da je v primeru regresije s fiksnimi učinki na deagregiranih podatkih nekoliko več neodvisnih spremenljivk, ki imajo statistično pomemben vpliv na odvisno spremenljivko takšnih, ki pripadajo nižjemu nivoju. Le ena spremenljivka, za katero se izkaže statistično pomemben vpliv, pripada višjemu nivoju. Gre za spremenljivko prisotnost pravil, ki tudi glede na večnivojsko analizo statistično pomembno vpliva na reflektivnost. Pri primerjavi rezultatov ugotovimo, da v tem primeru regresija podcenjuje njene učinke. Večnivojska analiza nam po drugi strani prikaže dve neodvisni spremenljivki s statistično pomembnim vplivom na reflektivnost. Poleg prisotnosti pravil ima po tej analizi statistično pomemben vpliv tudi popularnost foruma (rank alexa), ki pa po regresiji nima statistično pomembnega vpliva. Razlike se izkažejo tudi pri učinkih neodvisnih spremenljivk, ki pripadajo nižjemu nivoju. Teh je v primeru regresije nekoliko več kot pri večnivojski analizi. Po obeh metodah imajo statistično pomemben vpliv občutek pripadnosti, percepcija solidarnosti in zaznavanje formalnih pravil. Kot zanimivost se izkaže, da v vseh treh primerih regresija podcenjuje vplive teh spremenljivk. Glede na regresijsko analizo ima še ena neodvisna spremenljivka nižjega nivoja statistično pomemben vpliv na reflektivnost. To je spremenljivka zaznavanje neformalnih sankcij, ki pa se glede na privzeti interval zaupanja po večnivojski analizi ne izkaže za statistično pomembno.

Po pričakovanjih torej regresijska analiza nekoliko zanemara vplive neodvisnih spremenljivk višjega nivoja, saj ima po tej metodi le ena spremenljivka statistično pomemben vpliv na odvisno spremenljivko, medtem ko se pri večnivojski analizi dve spremenljivki izkažeta kot statistično pomembni. Po drugi strani se pri regresijski analizi nekoliko več neodvisnih spremenljivk nižjega nivoja izkaže kot statistično značilnih kot pri večnivojski analizi. Vsekakor se v konkretnem primeru izkaže, da strukturne spremenljivke ne igrajo bistvene vloge v pojasnjevanju reflektivnosti.

8. ZAKLJUČEK

Zanimanje za delovanje posameznikov v okvirih družbenih kontekstov, v katerih se le-ti nahajajo, je prisotno v večini socioloških teorij. V grobem torej lahko rečemo, da se sociologija zanima za delovanje posameznikov, ki pa je seveda pogojeno z družbo. Družbo lahko razumemo kot produkt človeške dejavnosti, pri čemer ne smemo zanemariti pomena medsebojnih interakcij. Seveda pa zagovorniki različnih teorij zagovarjajo tudi različne metode raziskovanja interakcij posameznikov. Med različnimi tezami predstavlja največjo ločnico preučevanje delovanja posameznikov po kvantitativni ali kvalitativni metodi. Ne glede na naklonjenost različnim načinom preučevanja, večina sociologov zagovarja tezo, da na življenje posameznikov ne vplivajo le njihove lastnosti, temveč tudi lastnosti družbenih skupin, v katerih se nahajajo.

Temeljna ideja, predstavljena v tem diplomskem delu, se torej nanaša na družbene skupine kot legitimne enote analize, kjer so lastnosti skupine ločene od lastnosti posameznih članov. Te lastnosti skupine, ki jih predstavljajo spremenljivke makro nivoja, lahko vplivajo na rezultate neodvisno od lastnosti posameznikov. Posledično nam torej večnivojska analiza pomaga razumeti posameznikova delovanja s pomočjo družbenih kontekstov. Do sedaj so znanstveniki pri tovrstnih konceptualnih problemih uporabljali tradicionalna statistična orodja, četudi so podatki in hipoteze bile večnivojske narave. Eden izmed takšnih pristopov je deagregacija informacij nivoja skupine na nivo posameznika, kjer so enote analize predstavljali posamezniki, kar je vodilo v dva večja problema. Vse nemodelirane kontekstualne informacije na koncu namreč pristanejo le v enem pojmu napake, kar je problematično, saj imajo posamezniki v isti skupini s tem že vnaprej predvidene korelirane napake, kar pa krši eno od predpostavk multiple regresije (Luke 2004). Naslednji problem, ki se nam v tem primeru pojavi pa je, da z neupoštevanjem skupin oziroma kontekstov, model predpostavlja, da so regresijski koeficienti enaki za vse skupine in s »tem sporoča, da se vsi procesi odvijajo na enak način v različnih kontekstih« (Duncan v Luke 2004).

Namen diplomskega dela je bila torej predstavitev večnivojske analize na teoretični ravni in preverjanje moči te metode na empirični ravni ter podati primerjavo večnivojske analize in regresije. Izkaže se, da so rezultati v primeru večnivojske analize bolj verodostojni, saj so povprečja na nivoju forumov v primeru regresije neustrezna. Kljub temu je v mojem konkretnem primeru zelo malo variabilnosti med forumi, kar pomeni, da strukturne spremenljivke ne igrajo bistvene vloge v pojasnjevanju reflektivnosti.

Z apliciranjem težavnosti večnivojske analize na problem, ki sem ga preučevala v svojem diplomskem delu je potrebno izpostaviti probleme, s katerimi sem se srečevala v analizah. Ugotovila sem namreč, da se v mojem primeru večnivojska analiza ni izkazala kot pomembna metoda z bistveno večjo pojasnjevalno močjo kot standardna multipla regresijska analiza. Glede na pojasnjen delež variance bi lahko trdila, da v bi v tem primeru lahko uporabili enostavno regresijsko analizo, saj med rezultati obeh analiz ni bistvenih razlik. Največja razlika se izkaže pri povprečjih strukturnih spremenljivk, pri kateri se večnivojska analiza izkaže za veliko bolj kredibilno kot standardna multipla regresijska analiza. To je posledica neupoštevanja skupin, saj so enote analize uporabniki forumov. Potrebno pa je poudariti, da model ne vključuje vseh dejavnikov ter lastnosti forumov, torej bi lahko spremenljivke nivoja dva razširili in s tem dobili boljše pojasnjevalne rezultate. Pomembno pa se je ustaviti tudi pri problematiki forumov. V tem primeru gre za spletne skupine, ki se od resničnih skupin razlikujejo v elementih, ki so lahko v večnivojski analizi ključni. V spletnih skupinah člani ne razvijejo nekega družbenega konsenza kot v primeru resničnih družbenih konstruktov. V spletnih forumih so prisotni različni komunikacijski in strukturni dejavniki, ki vplivajo na obstojnost in intenzivnost stikov med razpravljavci ter na njihov način izražanja mnenj. V tem kontekstu se v literaturi omenja predvsem vizualna anonimnost razpravljavcev in odsotnost normativno strukturiranih pravil (Petrič v Petrovčič 2005).

Ugotovila sem, da večnivojska analiza ni majhen zalogaj. Raziskovalec mora posedovati izpiljeno razumevanje družbenega sistema, ki ga preučuje, dobro razumevanje raziskovalnega problema, kjer se raziskovalec sooči z mnogimi vprašanji. Katere lastnosti posameznikov oziroma enot nivoja ena najboljše

opisujejo problem? Katere lastnosti skupine oziroma spremenljivke nivoja dva najboljše pojasnjuje problem? Kakšne so vzročno-posledične povezave? Kje je meja? V analizo je skorajda nemogoče vključiti vse pojasnjevalne spremenljivke, kar pomeni, da je pravilna izbira le-teh ključna za pridobitev veljavnih rezultatov. Eden izmed največjih izzivov je tudi zahteva po izpiljeni teoriji vzročnosti, ki povezuje spremenljivke makro in mikro nivoja in razlaga njihove povezave in interakcije med nivoji. Razviti je potrebno neodvisne razlage in preiti onstran uporabe večnivojske analize k statistični razlagi rezidualne variabilnosti med skupinami z upoštevanjem spremenljivk nivoja posameznika. Pri pomanjkanju le-tega tvegamo zmanjšanje pomembnosti večnivojske analize in jo skrhamo na nivo metode, ki ocenjuje variacijo med nepomembnimi skupinami ali med asociacijami z nepomembnimi spremenljivkami nivoja skupine. Odločilni element večnivojske analize je tudi natančna definicija ustrezne skupine in natančna definicija ustreznih spremenljivk nivoja skupine. Skupine niso poljubne vrste grupiranja posameznikov, temveč morajo predvsem imeti smisel pri razlagi rezultatov. Pri tem je ključna ideja, da spremenljivke nivoja skupin zagotavljajo bistvene informacije, ki jih v podatkih posameznikov ne moremo zajeti in s tem predstavljajo mere relevantnih konstruktov skupine. Za prihodnost nam torej še vedno ostaja veliko neopravljenega dela.

9. LITERATURA

- Diez, A V Roux (2002): *A glossary for multilevel analysis*. Dostopno na <http://jech.bmjournals.com> (19. julij 2006).
- Garson, David (2006): *Linear Mixed Models, Random Effects, Hierarchical Linear, Multilevel, Random Coefficients, and Repeated Measures Models*. Dostopno na <http://www2.chass.ncsu.edu/garson/pa765/multilevel.htm> (20. marec 2008).
- Goldstein, Harvey (1999): *Multilevel statistical models*. Dostopno na http://www.ats.ucla.edu/stat/examples/msm_goldstein/goldstein.pdf (15. marec 2008).
- Gupta, Vijay (2000): *Regression explained*. Dostopno na www.hsph.edu.vn/bmtinhoc/system/files/Statistics+-+SPSS+-+Regression.pdf (20. julij 2007).
- Hedeker, Don (1994): *Multilevel Analysis: An Applied Introduction*. Dostopno na http://www.uic.edu/classes/bstt/bstt513/ClustLS_SAS_SPSS.pdf (5. julij 2007).
- Hox, Joop (2002): *Multilevel Analysis, Techniques and Applications*. New Jersey: Lawrence Erlbaum Associates, Inc., Publishers.
- Hox, Joop (2002): *Multilevel Modeling: When and Why*. Amsterdam: TT-Publikaties.
- Hox, Joop (2002): *SPSS Mixed introduction*. Dostopno na <http://geocities.com/joophox/papers/spssmixedlab.pdf> (20. marec 2008).
- Jones, Kelvyn in Myles Gould (2005): *Developing multilevel models using MlwiN. A training manual*. Bristol: University of Bristol.

- Jones, Kelvyn (2006): *Multilevel modeling: practical applications*. Bristol: University of Bristol.
- Kalton, Graham in Vasja Vehovar (2001): *Vzorčenje v anketah*. Ljubljana: Fakulteta za družbene vede.
- Luke, A. Douglas (2004): *Multilevel modeling*. London: Sage publications, Inc.
- Osborne, W. Jason (2000): *Advantages of hierarchical linear modeling. Practical Assessment, Research & Evaluation*. Dostopno na <http://PAREonline.net/getvn.asp?v=7&n=1> (12. februar 2008).
- Osborne, W. Jason in Elain Waters (2002): *Four Assumptions of Multiple Regression That Researcher Should Always Test*. Dostopno na http://www4.ncsu.edu/~jwosbor2/otherfiles/PARE_2002_final.pdf (12. februar 2008).
- Petrič, Gregor (2007): Communicative community model of online social groups. *Let's play* 8, 17–20.
- Rasbash, Jon, Fiona Steele, William Brown in Bob Prosser (2004): *A User's Guide to MLwiN Version 2.0*. London: University of Bristol.
- Rasbash, Jon (2006): *Why use multilevel modelling?* Dostopno na http://www.cmm.bristol.ac.uk/learning-training/videos/jr-clioday_files/Default.htm#nopreload=1 (19. julij 2007).
- Raudenbush, W. Stephen in Anthony S. Bryk (2002): *Hierarchical Linear Models, Applications and data analysis methods*. California: Sage Publications, Inc.
- Schroeder, D. Larry, David L. Sjoquist in Paula E. Stephan (1986): *Understanding Regression Analysis. An Introductory Guide*. California: Sage Publications, Inc.
- Singer, D. Judith (1998): *Using SAS PROC MIXED to Fit Multilevel Models, Hierarchical Models, and Individual Growth Models*. Dostopno na

<http://gseweb.harvard.edu/%7Efaculty/singer/Papers/Using%20Proc%20Mixed.pdf> (19. julij 2007).

Snijders, A. B. Tom (2006): *Sampling*. Dostopno na <http://stat.gamma.rug.nl/sampling.pdf> (5. julij 2007).

Snijders, A. B. Tom (2006): *Multilevel Analysis*. Dostopno na <http://stat.gamma.rug.nl/MultilevelAnalysis.pdf> (5. julij 2007).

Sočan, Gregor (2006): *Regresijske analize : Večnivojsko modeliranje*. Dostopno na: http://193.2.70.110/Katedre/gradiva_podipl/magmethlm.ppt (5. julij 2007).

Viri:

Center for multilevel modelling (2008): *Learning and Training*. Dostopno na <http://www.cmm.bristol.ac.uk> (10. april 2008).

Scientific Software International (2008): *Hierarchical Linear and Nonlinear Modeling*. Dostopno na <http://www.ssicentral.com> (10. april 2008).

10. PRILOGE

PRILOGA A: REZULTATI VEČNIVOJSKE ANALIZE

Program: HLM 6 Hierarchical Linear and Nonlinear Modeling
Authors: Stephen Raudenbush, Tony Bryk, & Richard Congdon
Publisher: Scientific Software International, Inc. (c) 2000
techsupport@ssicentral.com
www.ssicentral.com

Module: HLM2.EXE (6.06.2857.2)
Date: 10 April 2008, Thursday
Time: 12: 2:21

SPECIFICATIONS FOR THIS HLM2 RUN

Problem Title: no title

The data source for this run = refleks4
The command file for this run = whlmtemp.hlm
Output file name = c:\\hlm2.txt
The maximum number of level-1 units = 1891
The maximum number of level-2 units = 48
The maximum number of iterations = 100
Method of estimation: restricted maximum likelihood

Weighting Specification

	Weighting?	Weight Variable Name	Normalized?
Level 1	no		
Level 2	no		
Precision	no		

The outcome variable is REFLEKSI

The model specified for the fixed effects was:

Level-1 Coefficients		Level-2 Predictors
INTRCPT1,	B0	INTRCPT2, G00
\$		RAGE_MEA, G01
\$		RANK_ALE, G02
		TIP1_MEA, G03
\$		PRAVILA, G04
\$		IDENTIFI, G05
##% OBCUTEKP slope,	B1	INTRCPT2, G10
##% PERCEPCI slope,	B2	INTRCPT2, G20
##% PREZEMAN slope,	B3	INTRCPT2, G30
##% FORMALNE slope,	B4	INTRCPT2, G40
##% V8_A slope,	B5	INTRCPT2, G50

#% NEFORMAL slope, B6 INTRCPT2, G60
#% POZITIVN slope, B7 INTRCPT2, G70
#% IDENTIFI slope, B8 INTRCPT2, G80
#% KOHERENT slope, B9 INTRCPT2, G90

'#' - The residual parameter variance for this level-1 coefficient has been set to zero.

'%' - This level-1 predictor has been centered around its grand mean.

'\$' - This level-2 predictor has been centered around its grand mean.

The model specified for the covariance components was:

Sigma squared (constant across level-2 units)

Tau dimensions
INTRCPT1

Summary of the model specified (in equation format)

Level-1 Model

$$Y = B0 + B1*(OBCUTEKP) + B2*(PERCEPCI) + B3*(PREZEMAN) + B4*(FORMALNE) + B5*(V8_A) + B6*(NEFORMAL) + B7*(POZITIVN) + B8*(IDENTIFI) + B9*(KOHERENT) + R$$

Level-2 Model

$$B0 = G00 + G01*(RAGE_MEA) + G02*(RANK_ALE) + G03*(TIP1_MEA) + G04*(PRAVILA) + G05*(IDENTIFI) + U0$$

B1 = G10
B2 = G20
B3 = G30
B4 = G40
B5 = G50
B6 = G60
B7 = G70
B8 = G80
B9 = G90

Run-time deletion has reduced the number of level-1 records to 1658

Iterations stopped due to small change in likelihood function

***** ITERATION 41 *****

Sigma_squared = 0.45354

Tau
INTRCPT1,B0 0.00892

Tau (as correlations)
INTRCPT1,B0 1.000

Random level-1 coefficient Reliability estimate

 INTRCPT1, B0 0.294

The value of the likelihood function at iteration 41 = -1.751037E+003
 The outcome variable is REFLEKSI

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard Error	T-ratio	Approx. d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	3.791068	0.091642	41.368	42	0.000
RAGE_MEA, G01	-0.008403	0.014983	-0.561	42	0.577
RANK_ALE, G02	0.000046	0.000027	1.718	42	0.093
TIP1_MEA, G03	-0.046113	0.061914	-0.745	42	0.461
PRAVILA, G04	0.061398	0.027357	2.244	42	0.030
IDENTIFI, G05	0.028679	0.034290	0.836	42	0.408
For OBCUTEKP slope, B1					
INTRCPT2, G10	0.162302	0.029327	5.534	1643	0.000
For PERCEPCI slope, B2					
INTRCPT2, G20	0.199766	0.033881	5.896	1643	0.000
For PREZEMAN slope, B3					
INTRCPT2, G30	0.006883	0.009617	0.716	1643	0.474
For FORMALNE slope, B4					
INTRCPT2, G40	0.098320	0.024475	4.017	1643	0.000
For V8_A slope, B5					
INTRCPT2, G50	0.011765	0.015389	0.765	1643	0.445
For NEFORMAL slope, B6					
INTRCPT2, G60	-0.037623	0.016231	-2.318	1643	0.021
For POZITIVN slope, B7					
INTRCPT2, G70	0.008389	0.014439	0.581	1643	0.561
For IDENTIFI slope, B8					
INTRCPT2, G80	-0.001053	0.006982	-0.151	1643	0.880
For KOHERENT slope, B9					
INTRCPT2, G90	0.000548	0.009479	0.058	1643	0.954

The outcome variable is REFLEKSI

Final estimation of fixed effects
 (with robust standard errors)

Fixed Effect	Coefficient	Standard Error	T-ratio	Approx. d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	3.791068	0.084712	44.753	42	0.000
RAGE_MEA, G01	-0.008403	0.012419	-0.677	42	0.502
RANK_ALE, G02	0.000046	0.000020	2.254	42	0.029
TIP1_MEA, G03	-0.046113	0.054033	-0.853	42	0.398
PRAVILA, G04	0.061398	0.028253	2.173	42	0.035
IDENTIFI, G05	0.028679	0.026118	1.098	42	0.279
For OBCUTEKP slope, B1					
INTRCPT2, G10	0.162302	0.032034	5.067	1643	0.000
For PERCEPCI slope, B2					
INTRCPT2, G20	0.199766	0.039625	5.041	1643	0.000
For PREZEMAN slope, B3					
INTRCPT2, G30	0.006883	0.010173	0.677	1643	0.499

For FORMALNE slope, B4 INTRCPT2, G40	0.098320	0.020708	4.748	1643	0.000
For V8_A slope, B5 INTRCPT2, G50	0.011765	0.016204	0.726	1643	0.468
For NEFORMAL slope, B6 INTRCPT2, G60	-0.037623	0.019441	-1.935	1643	0.053
For POZITIVN slope, B7 INTRCPT2, G70	0.008389	0.013742	0.610	1643	0.541
For IDENTIFI slope, B8 INTRCPT2, G80	-0.001053	0.008048	-0.131	1643	0.896
For KOHERENT slope, B9 INTRCPT2, G90	0.000548	0.009352	0.059	1643	0.954

Final estimation of variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, U0	0.09445	0.00892	42	64.73916	0.014
level-1, R	0.67345	0.45354			

Statistics for current covariance components model

Deviance = 3502.074132
Number of estimated parameters = 2

PRILOGA B – REZULTATI REGRESIJE S FIKSNIMI UČINKI

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
ZFORUM Zscore(FORUM)	1895	-.80697	3.26516	.0000000	1.0000000
ZRAge Zscore: Dejanska Starost foruma	1895	-1.84135	1.69980	.0000000	1.0000000
ZRank_Alexa Zscore (Rank_Alexa)	1895	-.42605	10.60753	.0000000	1.0000000
ZTIP1 Zscore(TIP1)	1895	-1.09791	.91034	.0000000	1.0000000
ZPravila Zscore: Stopnja formalizacije pravil	1895	-2.62542	1.89028	.0000000	1.0000000
ZIdentifikacija Zscore (Identifikacija)	1895	-.58803	2.08419	.0000000	1.0000000
ZRefleksivnost Zscore (Refleksivnost)	1849	-3.72171	1.78510	.0000000	1.0000000
ZObcutekPripadnosti Zscore (ObcutekPripadnosti)	1824	-2.70319	2.86761	.0000000	1.0000000
ZPercepcijaSolidarnosti Zscore (PercepcijaSolidarnosti)	1821	-4.85949	1.69724	.0000000	1.0000000
ZPrezemanjeOffline Zscore (PrezemanjeOffline)	1885	-.61417	3.96628	.0000000	1.0000000
ZFormalnePravila Zscore (FormalnePravila)	1852	-3.14306	.73843	.0000000	1.0000000
ZFormalneSankcije Zscore (FormalneSankcije)	1818	-2.42500	3.84908	.0000000	1.0000000
ZNeformalneSankcije Zscore (NeformalneSankcije)	1811	-1.84649	4.30881	.0000000	1.0000000
ZPozitivneSankcije Zscore(PozitivneSankcije)	1809	-2.33296	3.44155	.0000000	1.0000000
ZIdentifikabilnost Zscore (Identifikabilnost)	1796	-1.22069	3.97534	.0000000	1.0000000
ZKoherentnost Samopredstavitve Zscore (Koherentnost Samopredstavitve)	1815	-4.33838	.89327	.0000000	1.0000000
Valid N (listwise)	1658				

Variables Entered/Removed^b

Model	Variables Entered	Variables Removed	Method
1	ZKoherentnostSamopredstavitev Zscore (Koherentnost Samopredstavitev), ZPrezemanje Offline Zscore (Prezemanje Offline), ZIdentifikacija Zscore (Identifikacija), ZPozitivne Sankcije Zscore (Pozitivne Sankcije), ZRAge Zscore: Dejanska Starost foruma, ZPravila Zscore: Stopnja formalizacije pravil, ZPercepcija Solidarnosti Zscore (Percepcija Solidarnosti), ZFormalne Pravila Zscore (Formalne Pravila), ZRank_Alexa Zscore (Rank_Alexa), ZIdentifikabilnost Zscore (Identifikabilnost), ZFormalne Sankcije Zscore (Formalne Sankcije), ZObcutek Pripadnosti Zscore (Obcutek Pripadnosti), ZNeformalne Sankcije Zscore (Neformalne Sankcije), ZTIP1 Zscore (TIP1) ^a		Enter

a. All requested variables entered.

b. Dependent Variable: ZRefleksivnost Zscore(Refleksivnost)

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.330 ^a	.109	.101	.93305743

a. Predictors: (Constant), ZKoherentnostSamopredstavitev Zscore(KoherentnostSamopredstavitev), ZPrezemanjeOffline Zscore(PrezemanjeOffline), ZIdentifikacija Zscore (Identifikacija), ZPozitivneSankcije Zscore(PozitivneSankcije), ZRAge Zscore: Dejanska Starost foruma, ZPravila Zscore: Stopnja formalizacije pravil, ZPercepcijaSolidarnosti Zscore(PercepcijaSolidarnosti), ZFormalnePravila Zscore(FormalnePravila), ZRank_Alexa Zscore(Rank_Alexa), ZIdentifikabilnost Zscore(Identifikabilnost), ZFormalneSankcije Zscore(FormalneSankcije), ZObcutekPripadnosti Zscore(ObcutekPripadnosti), ZNeformalneSankcije Zscore(NeformalneSankcije), ZTIP1 Zscore(TIP1)

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	174.857	14	12.490	14.346	.000 ^a
	Residual	1430.389	1643	.871		
	Total	1605.246	1657			

a. Predictors: (Constant), ZKoherentnostSamopredstavitve Zscore(KoherentnostSamopredstavitve), ZPrezemanjeOffline Zscore(PrezemanjeOffline), ZIdentifikacija Zscore (Identifikacija), ZPozitivneSankcije Zscore(PozitivneSankcije), ZRAge Zscore: Dejanska Starost foruma, ZPravila Zscore: Stopnja formalizacije pravil, ZPercepcijaSolidarnosti Zscore(PercepcijaSolidarnosti), ZFormalnePravila Zscore(FormalnePravila), ZRank_Alexa Zscore(Rank_Alexa), ZIdentifikabilnost Zscore(Identifikabilnost), ZFormalneSankcije Zscore(FormalneSankcije), ZObcutekPripadnosti Zscore(ObcutekPripadnosti), ZNeformalneSankcije Zscore(NeformalneSankcije), ZTIP1 Zscore(TIP1)

b. Dependent Variable: ZRefleksivnost Zscore(Refleksivnost)

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.005	.023		.228	.820
	ZRAge Zscore: Dejanska Starost foruma	-.020	.035	-.020	-.565	.572
	ZRank_Alexa Zscore (Rank_Alexa)	.045	.025	.047	1.771	.077
	ZTIP1 Zscore(TIP1)	-.006	.034	-.006	-.187	.852
	ZPravila Zscore: Stopnja formalizacije pravil	.052	.024	.053	2.165	.031
	ZIdentifikacija Zscore (Identifikacija)	.029	.024	.029	1.225	.221
	ZObcutekPripadnosti Zscore (ObcutekPripadnosti)	.159	.029	.161	5.532	.000
	ZPercepcijaSolidarnosti Zscore (PercepcijaSolidarnosti)	.158	.028	.156	5.611	.000
	ZPrezemanjeOffline Zscore (PrezemanjeOffline)	.013	.023	.013	.574	.566
	ZFormalnePravila Zscore (FormalnePravila)	.092	.026	.093	3.581	.000
	ZFormalneSankcije Zscore (FormalneSankcije)	.018	.027	.019	.679	.497
	ZNeformalneSankcije Zscore (NeformalneSankcije)	-.066	.029	-.068	-2.282	.023
	ZPozitivneSankcije Zscore(PozitivneSankcije)	.016	.028	.017	.597	.551
	ZIdentifikabilnost Zscore (Identifikabilnost)	-.013	.025	-.013	-.524	.601
	ZKoherentnost Samopredstavitve Zscore (Koherentnost Samopredstavitve)	.009	.024	.009	.364	.716

a. Dependent Variable: ZRefleksivnost Zscore(Refleksivnost)