Isplati li se transparentnost? Utjecaj proračunske transparentnosti na fiskalne rezultate hrvatskih općina i gradova

Marijan, Lucija

SVEUČILIŠTE U ZAGREBU

PRIRODOSLOVNO–MATEMATIČKI FAKULTET

MATEMATIČKI ODSJEK

Lucija Marijan

ISPLATI LI SE TRANSPARENTNOST? UTJECAJ PRORAČUNSKE TRANSPARENTNOSTI NA FISKALNE REZULTATE HRVATSKIH OPĆINA I GRADOVA

Diplomski rad

Voditelj rada:
dr. sc. Katarina Ott

Zagreb, 2022.

Ovaj diplomski rad obranjen je dana ____________________ pred ispitnim povjerenstvom u sastavu:

1. ____________________________, predsjednik
2. ____________________________, član
3. ____________________________, član

Povjerenstvo je rad ocijenilo ocjenom _______________.

Potpisi članova povjerenstva:

1. ____________________________
2. ____________________________
3. ____________________________

Sadržaj

Uvod

Proračunska transparentnost pojam je koji sve češće čujemo u medijima te je sve popularniji pristup približavanja građanima i zadobivanja njihovog povjerenja. Međunarodne organizacije kao što su OECD (Organisation for Economic Co-operation and Development) i MMF (Međunarodni monetarni fond) iniciraju veću transparentnost. U Hrvatskoj, analizu proračunske transparentnosti jednom godišnje provodi Institut za javne financije na razini lokalnih i područnih samouprava. Cilj ovog rada je utvrditi utječe li proračunska transparentnost na fiskalne rezultate hrvatskih općina i gradova.

Institut za javne financije od 2015. jednom godišnje objavljuje analizu proračunske transparentnosti, koja ovisno o broju ključnih proračunskih dokumenata objavljenih na službenim mrežnim stranicama lokalnih samouprava, iznosi od 0 do 5. Nezavisna varijabla u ovom radu je proračunska transparentnost, a zavisne varijable su rashod po stanovniku, prihod po stanovniku, suficit po stanovniku i javni dug po stanovniku 2020. godine.

Na temelju radova [1], [2], [3], [4], [5], [18] postavljene su četiri hipoteze. Prva hipoteza tvrdi da proračunska transparentnost doprinosi povećanju proračunskog rashoda. Druga hipoteza govori da proračunska transparentnost doprinosi povećanju proračunskog prihoda. Treća hipoteza testira vezu između proračunske transparentnosti i proračunskog suficita, odnosno da li veća transparentnost doprinosi ostvarenju proračunskog suficita. Četvrta hipoteza tvrdi da proračunska transparentnost doprinosi smanjenju javnog duga.

Modeli kojima će se testirati hipoteze izabrani su prema podacima. Hipoteze će se testirati jednostavnom linearnom regresijom. Značajnost transparentnosti može ovisiti i o drugim izvorima, stoga su uvedene kontrolne varijable postotak korisnika interneta, stopa nezaposlenosti i prosječni dohodak. Kontrolne varijable uvedene su na temelju [14], [15] i [21]. Modelom višestruke linearne regresije, također će se testirati hipoteze. Analiza podataka bit će provedena u programskom jeziku R, posebno na razini gradova, a posebno na razini općina. Testiranjem hipoteza utvrđeno je da ne postoji statistički značajna veza između proračunske transparentnosti i fiskalnih rezultata hrvatskih općina i gradova, odnosno sve hipoteze su odbačene. Uvođenjem kontrolnih varijabli utjecaj transparentnosti ostaje statistički neznačajan za sve zavisne varijable.

Rad se sastoji od 5 poglavlja. U prvom poglavlju dana je definicija proračunske transparentnosti te su postavljene hipoteze koje će se testirati. U drugom poglavlju detaljnije su opisane varijable korištene radu. U trećem poglavlju navedena je metodologija, odnosno modeli koji će se koristiti u radu, te njihova matematička podloga. U četvrtom poglavlju testirane su hipoteze te je navedeno jesu li one prihvaćene ili odbačene. U petom poglavlju naveden je zaključak rada nakon kojeg slijedi sažetak te životopis.

Poglavlje 1

Proračunska transparentnost

1.1 Pojam proračunske transparentnosti

Proračunska transparentnost pojam je koji se sve češće koristi u svakodnevnom govoru, a ipak nema jedinstvenu definiciju. Najčešće citirana je ona od Kopitsa i Craiga [12]. Prema njima proračunska transparentnost je otvorenost prema javnosti u pogledu strukture i funkcija vlade, ciljeva fiskalne politike, računa javnog sektora te projekcija za budućnost. To uključuje siguran pristup pouzdanim, jasnim, razumljivim te međunarodno usporedivim informacijama o aktivnostima vlade kako bi birači i financijska tržišta mogla precizno procijeniti financijsku poziciju države te stvarne troškove i koristi vladinih aktivnosti, uključujući njihove sadašnje i buduće ekonomske i društvene posljedice.

Prema Institutu za javne financije [20], proračunska transparentnost podrazumijeva uvid u potpune, točne, pravovremene i razumljive informacije. Na temelju njih, građani se mogu angažirati i, između ostalog, pokušati utjecati na efikasnost prikupljanja i trošenja proračunskih sredstava, odgovornosti Vlade i vlasti lokalnih jedinica, kao i na smanjenje mogućih koruptivnih radnji.

Proračunska transparentnost sve je popularniji pristup približavanja građanima i zadobivanja njihovog povjerenja. Uvidom u stanje proračuna građani mogu saznati koliko kvalitetno se upravlja njihovom lokalnom jedinicom. No, nažalost, građani, ili nemaju interesa i vremena, ili su im podaci nerazumljivi. Stoga je promicanje važnosti transparentnosti jedan od načina napretka društva, a samim time i napretka javne vlasti.

1.2 Hipoteze

Ovim radom testirat će se utjecaj proračunske transparentnosti na proračunski rashod po stanovniku, proračunski prihod po stanovniku, proračunski suficit po stanovniku te javni dug po stanovniku. Dodatno, uz transparentnost, testirat će se imaju li stopa nezaposlenosti, prosječni dohodak i postotak korisnika interneta utjecaj na gore navedene proračunske varijable. U ovom odjeljku navedene su hipoteze koje su donesene na temelju [1], [2], [3], [4], [5], [18].
Hipoteze će se testirati posebno po gradovima i posebno po općinama, te dodatno s kontrolnim varijablama.

1.2.1 Hipoteza 1

Hipoteza 1: Veća transparentnost doprinosi povećanju proračunskog rashoda.

Iako bi bilo za očekivati da veća transparentnost tjera političare na odgovorniju fiskalnu politiku, pa time i smanjenje rashoda, u istraživanjima se pokazalo da to nije tako. Prema Alt, Lassen i Skilling [4] veća transparentnost povezana je s većim proračunskim rashodom. Jedan od razloga za takav rezultat može se pronaći u tome da većom transparentnošću političari više ne mogu podcijeniti visinu rashoda u njihovim prijedlozima proračuna.

1.2.2 Hipoteza 2

Hipoteza 2: Veća proračunska transparentnost doprinosi povećanju proračunskog prihoda.

Veća transparentnost rezultira većim povjerenjem prema političarima, što dovodi do toga da su im građani spremni povjeriti više sredstava. Odnosno, većom transparentnošću građani su spremni plaćati veće poreze. Prema Alt, Lassen i Skilling [4] veća transparentnost povezana je s većim proračunskim prihodom.

1.2.3 Hipoteza 3

Hipoteza 3: Veća proračunska transparentnost doprinosi ostvarenju proračunskog suficita.

Veća transparentnost onemogućuje političarima korištenje fiskalnog deficita za ostvarivanje vlastitih ciljeva. Prema Benito i Bastida [5], veća proračunska transparentnost doprinosi ostvarenju proračunskog suficita.

1.2.4 Hipoteza 4

Hipoteza 4: Veća proračunska transparentnost doprinosi smanjenju duga lokalnih jedinica.

Veća transparentnost tjera političare na odgovorniju fiskalnu politiku, što utječe na smanjenje akumulacije javnog duga. Prema Alt i Lassen [3] veća proračunska transparentnost doprinosi smanjenju duga lokalnih jedinica.

U prvom poglavlju dana je definicija transparentnost te njena važnost koja je motiv za mnoge radove, kao i za ovaj. Navedene su hipoteze koje će se testirati, kao i prethodni radovi na kojima su te hipoteze donesene. U navedenim radovima dokazana je statistički značajna veza između transparentnosti i proračunskog rashoda, prihoda, suficita i javnog duga po stanovniku. U daljnjem dijelu rada utvrdit će se vrijede li postavljene hipoteze i za hrvatske gradove i općine. Ako se pokaže da postoji statistički značajna veza između proračunske transparentnosti i proračunskog rashoda, prihoda, suficita i javnog duga, to bi trebao biti dodatan motiv za njezino promicanje.

Poglavlje 2

Podaci

U ovom odjeljku opisani su podaci koji će se koristiti u testiranju hipoteza navedenih u prethodnom poglavlju te koji su izvori tih podataka.

2.1 Nezavisna varijabla

Nezavisna varijabla u ovom radu je proračunska transparentnost koju već duži niz godina analizira Institut za javne financije. Proračunska transparentnost u Hrvatskoj mjeri se brojem ključnih proračunskih dokumenata objavljenih na službenim mrežnim stranicama lokalnih samouprava. Dokumenti koji se gledaju su prijedlog proračuna, izglasani proračun, polugodišnje izvješće o izvršenju proračuna, godišnje izvješće o izvršenju proračuna i proračunski vodič za građane. Takvim mjerenjem razina proračunske transparentnosti može iznositi od 0 do 5. Podaci o proračunskoj transparentnosti jedinica lokalnih samouprava uzeti su za 2020. godinu sa stranice Instituta za javne financije [6].

2.2 Zavisne varijable

Zavisne varijable u ovom radu su proračunski prihod po stanovniku, proračunski rashod po stanovniku, proračunski suficit po stanovniku i javni dug po stanovniku hrvatskih općina i gradova.

Prihod jedinica lokalnih samouprava čine porezi, prihodi od pomoći i prihodi od vlastitih i namjenskih prihoda. Prihod po stanovniku računa se kao omjer prihoda i broja stanovnika jedinice lokalne samouprave.

Rashodi jedinica lokalnih samouprava odražavaju način upotrebe proračunskih prihoda radi zadovoljenja javnih interesa. Rashod po stanovniku računa se kao omjer rashoda i broja stanovnika jedinice lokalne samouprave.

Proračunski suficit višak je prihoda nad rashodima proračuna u određenom vremenskom razdoblju. Suficit po stanovniku računa se kao omjer ostvarenog suficita i broja stanovnika jedinice lokalne samouprave.

Javni dug ukupna je zaduženost jedinice lokalne samouprave prema domaćim i inozemnim vjerovnicima. Javni dug po stanovniku računa se kao omjer javnog duga i broja stanovnika jedinice lokalne samouprave.

Podaci za sve zavisne varijable uzeti su za 2020. godinu sa stranica Ministarstva financija RH [13].

2.3 Kontrolne varijable

Kontrolne varijable uvrštene su s proračunskom transparentnošću u model višestruke linearne regresije radi dobivanja preciznijeg utjecaja transparentnosti na zavisne varijable. U ovom radu kontrolne varijable su stopa nezaposlenosti, prosječni dohodak po stanovniku i postotak korisnika interneta te su uzete na temelju radova [14] , [15] i [21].

Podatak o postotku korisnika interneta izračunat je u sklopu popisa stanovništva za 2011. godinu te je preuzet sa stranice Državnog zavoda za statistiku [8]. Računa se kao omjer osoba koje imaju sljedeća znanja: pronaći internetsku stranicu putem tražilice, dolazak na internetske stranice upisom adrese ili putem veze, korištenjem mapom Favorites, postava početne stranice, otvaranje stranice u novom prozoru itd., i broja stanovnika koji žive na području jedinice lokalne samouprave.

Stopa nezaposlenosti uzeta je iz podloge za izračun indeksa razvijenosti za 2017. godinu [16]. Ona se računa kao omjer nezaposlenih i zbroja svih zaposlenih te nezaposlenih osoba na području jedinice lokalne samouprave.

Prosječni dohodak uzet je iz podloge za izračun indeksa razvijenosti za 2017. godinu [16]. On se računa kao omjer ukupnog iznosa dohotka kojega su tijekom jednog poreznog razdoblja ostvarili porezni obveznici, fizičke osobe s prebivalištem ili uobičajenim boravištem na području jedinice lokalne samouprave, i broja stanovnika koji žive na području te jedinice.

2.4 Deskriptivna statistika

U ovom odjeljku bit će opisani podaci radi njihovog boljeg razumijevanja. Varijable su podijeljene na gradove i općine te je za njih posebno napravljena deskriptivna statistika.

2.4.1 Deskriptivna statistika transparentnosti

U tablici 2.1 prikazana je frekvencija i relativna frekvencija transparentnosti proračuna gradova, dok su u tablici 2.2 prikazana statistička obilježja transparentnosti gradova.

Razina transparentnosti	Frekvencija	Relativna frekvencija
0	0	0
1	1	0,008
2	4	0,031
3	10	0,078
4	27	0,211
5	86	0,672

Tablica 2.1: Frekvencija i relativna frekvencija proračunske transparentnosti gradova

	Transparentnost
Minimum	1
Prvi kvartil	4
Medijan	5
Treći kvartil	5
Maksimum	5
Aritmetička sredina	4,508

Tablica 2.2: Statistička obilježja transparentnosti gradova

Iz tablice 2.1 može se iščitati da niti jedan grad nema transparentnost 0, dok 86 gradova ima transparentnost 5, odnosno čak 67% gradova objavljuje sve potrebne dokumente na svojima internetskim stranicama. Općenito, frekvencija raste kako raste razina transparentnosti gradova.

U tablici 2.3 prikazana je frekvencija i relativna frekvencija transparentnosti proračuna općina, dok su u tablici 2.4 prikazana statistička obilježja transparentnosti općina. Kao i kod gradova, frekvencija raste kako raste razina transparentnosti općina.

Razina transparentnosti	Frekvencija	Relativna frekvencija
0	7	0,016
1	17	0,040
2	33	0,077
3	58	0,136
4	111	0.259
5	202	0,472

Tablica 2.3: Frekvencija i relativna frekvencija proračunske transparentnosti općina

	Transparentnost
Minimum	0
Prvi kvartil	3
Medijan	4
Treći kvartil	5
Maksimum	5
Aritmetička sredina	3,998

Tablica 2.4: Statistička obilježja transparentnosti općina

Na temelju gornjih tablica može se zaključiti da općine zaostaju za gradovima što se tiče transparentnosti. Dok je za gradove medijan transparentnosti jednak 5, za općine je on jednak 4. Dodatno, niti jedan dokument na svojim internetskim stranicama ne objavljuje 7 općina, od kojih se 3 nalaze u Zadarskoj županiji.
Treba napomenuti da su lokalne samouprave zakonom obvezne objavljivati Izvještaj o godišnjem i polugodišnjem proračunu te Izglasani proračun, stoga iz gornjih tablica slijedi da najmanje 4% gradova i 13% općina ne provode zakonske odredbe.
Niža transparentnost kod općina u odnosu na gradove može biti zbog toga što su ljudi u gradovima uključeniji u političke procese pa samim time i u efikasnost upravljanja lokalnim samoupravama.

2.4.2 Deskriptivna statistika rashoda

U tablici 2.5 prikazana je aritmetička sredina rashoda po stanovniku u odnosu na pojedinu razinu transparentnosti proračuna gradova i općina.

Razina transparentnosti	Gradovi	Općine
0	0	6.931
1	4.068	5.419
2	7.351	5.021
3	7.461	5.973
4	6.586	5.674
5	6.151	5.501

Tablica 2.5: Aritmetička sredina rashoda po stanovniku u odnosu na razinu proračunske transparentnosti

U tablici 2.6 prikazana su statistička obilježja rashoda po stanovniku.

	Gradovi	Općine
Minimum	2.646	1.857
Prvi kvartil	4.737	3.613
Medijan	5.568	4.760
Treći kvartil	7.664	6.821
Maksimum	18.549	26.802
Aritmetička sredina	6.366	5.593

Tablica 2.6: Statistička obilježja rashoda po stanovniku

Iz tablice 2.5 može se iščitati da najveći prosječni rashod po stanovniku imaju gradovi s razinom transparentnosti 3, dok najveći prosječni rashod po stanovniku imaju općine s razinom transparentnosti 0. Gradovi s transparentnosti 3 imaju najveći prosječni rashod po stanovniku jer u tu skupinu spada samo 10 gradova, od kojih jedan ima najveći rashod po stanovniku od svih gradova u RH. Razlog najvećih prosječnih rashoda po stanovniku u općinama s transparentnosti 0 može biti to što političari troše više javnih sredstava budući da građanima nisu lako dostupne informacije u koje svrhe se ta sredstva upotrebljavaju. Kako se aritmetička sredina rashoda po stanovniku ne povećava s većom razinom transparentnosti možemo očekivati da će prva hipoteza biti odbačena.
Iz Tablice 2.6 se može iščitati da najveći rashod po stanovniku pripada općini, ali da gradovi imaju veći prosječni rashod po stanovniku od općina.

2.4.3 Deskriptivna statistika prihoda

U tablici 2.7 prikazana je aritmetička sredina prihoda po stanovniku u odnosu na pojedinu razinu transparentnosti proračuna gradova i općina.

Razina transparentnosti	Gradovi	Općine
0	0	7.789
1	3.955	5.713
2	6.286	5.199
3	7.485	6.357
4	6.192	6.144
5	6.474	5.798

Tablica 2.7: Aritmetička sredina prihoda po stanovniku u odnosu na razinu proračunske transparentnosti

Iz tablice 2.7 može se iščitati da najveći prosječni prihod po stanovniku imaju gradovi s razinom transparentnosti 3, dok najveći prosječni prihod po stanovniku imaju općine s razinom transparentnosti 0. Iz navedene tablice može se primijetiti da takva raspodjela prihoda po stanovniku u odnosu na pojedinu razinu proračunske transparentnosti ide kontra drugoj hipotezi.

U tablici 2.8 prikaza su statistička obilježja prihoda po stanovniku hrvatskih općina i gradova.

	Gradovi	Općine
Minimum	2.934	2.147
Prvi kvartil	4.939	3.767
Medijan	5.938	5.126
Treći kvartil	7.873	6.938
Maksimum	16.346	27.035
Aritmetička sredina	6.468	5.947

Tablica 2.8: Statistička obilježja prihoda po stanovniku

Iz tablice 2.8 vidi se da gradovi imaju veći prosječni prihod po stanovniku od općina, dok najveći prihod po stanovniku od svih lokalnih samouprava pripada općini. Veći prosječni prihod po stanovniku u gradovima u odnosu na općine može biti posljedica većeg prosječnog dohotka stanovnika u gradovima.

2.4.4 Deskriptivna statistika suficita

Od 128 gradova, 81 ostvaruje proračunski suficit, dok od 428 općina, njih 290 ostvaruje proračunski suficit.

U tablici 2.9 prikazana je aritmetička sredina suficita u odnosu na pojedinu razinu transparentnosti proračuna.

Razina transparentnosti	**Gradovi**	**Općine**
0	0	1.260
1	0	815
2	84	551
3	1.176	815
4	291	966
5	679	793

Tablica 2.9: Aritmetička sredina suficita po stanovniku u odnosu na razinu proračunske transparentnosti

U tablici 2.10 prikazana su statistička obilježja suficita.

	Gradovi	**Općine**
Minimum	8	3
Prvi kvartil	197	271
Medijan	434	548
Treći kvartil	852	1.104
Maksimum	5.052	5.615
Aritmetička sredina	657	830

Tablica 2.10: Statistička obilježja suficita po stanovniku

Iz tablice 2.9 može se iščitati da najveći prosječni suficit po stanovniku imaju gradovi s razinom transparentnosti 3, dok najveći prosječni suficit po stanovniku imaju općine s razinom transparentnosti 0. Prema potpoglavljima 2.4.3 i 2.4.4 vidljivo je da općine s transparentnosti 0 imaju najveće rashode, ali još veće prihode, stoga je i za očekivati da će one ostvarivati najveći proračunski suficit. Također, iz tablice je vidljivo da gradovi s transparentnosti 1 ne ostvaruju proračunski suficit. Prema aritmetičkoj sredini suficita u odnosu na pojedinu razinu transparentnosti možemo očekivati da će treća hipoteza biti odbačena.
Veći prosječni suficit ostvaruju gradovi, što je vidljivo iz tablice 2.10. Iz iste tablice slijedi da najveći proračunski suficit pripada općini.

2.4.5 Deskriptivna statistika javnog duga

Od 128 gradova, njih 108 imaju javni dug, dok od 428 općina, njih 261 imaju javni dug.

U tablici 2.11 prikazana je aritmetička sredina javnog duga po stanovniku u odnosu na pojedinu razinu transparentnosti proračuna.

Razina transparentnosti	Gradovi	Općine
0	0	1.579
1	0	1.101
2	1.536	1.623
3	2.982	1.113
4	1.001	1.426
5	1.552	1.338

Tablica 2.11: Aritmetička sredina javnog duga po stanovniku u odnosu na razinu proračunske transparentnosti

U tablici 2.12 prikazana su statistička obilježja javnog duga.

	Gradovi	Općine
Minimum	10	0,002
Prvi kvartil	449	240
Medijan	1.008	787
Treći kvartil	1.913	1.704
Maksimum	7.802	9.809
Aritmetička sredina	1.550	1.347

Tablica 2.12: Statistička obilježja javnog duga po stanovniku

Iz tablice 2.11 može se iščitati da najveći prosječni javni dug stanovniku imaju gradovi s razinom transparentnosti 3, dok najveći prosječni javni dug po stanovniku imaju općine s razinom transparentnosti 2. Iz tablice je također vidljivo da gradovi s transparentnosti 1 nemaju javni dug. Kako se aritmetička sredina javnog duga po stanovniku ne povećava s većom razinom transparentnosti možemo očekivati da će četvrta hipoteza biti odbačena. Iz tablice 2.12 slijedi da gradovi imaju veći prosječni javni dug po stanovniku, ali da najveći javni dug po stanovniku pripada općini.

2.4.6 Deskriptivna statistika kontrolnih varijabli

U ovom odjeljku bit će prikaza deskriptivna statistika postotka korisnika interneta (u tablicama označeno s I), stope nezaposlenosti (u tablicama označeno N) i prosječnog dohotka (u tablicama označeno D).

U tablici 2.13 prikaza su statistička obilježja kontrolnih varijabli gradova, dok su u tablici 2.14 statistička obilježja kontrolnih varijabli općina.

	I	**N**	**D**
Minimum	0,395	0,042	18.630
Prvi kvartil	0,501	0,099	24.675
Medijan	0,541	0,150	28.238
Treći kvartil	0,571	0,204	32.153
Maksimum	0,707	0.327	44.733
Aritmetička sredina	0,537	0,156	28.759

Tablica 2.13: Statistička obilježja kontrolnih varijabli gradova

	I	**N**	**D**
Minimum	0,178	0,035	8.948
Prvi kvartil	0,423	0,105	18.619
Medijan	0,465	0,165	22.133
Treći kvartil	0,508	0,257	26.825
Maksimum	0,680	0,524	39.272
Aritmetička sredina	0,465	0,186	22.990

Tablica 2.14: Statistička obilježja kontrolnih varijabli općina

Usporedbom tablica 2.13 i 2.14 slijedi da gradovi imaju veći prosječni postotak korisnika interneta što je i za očekivati budući da su gradovi urbaniji te imaju bolju telekomunikacijsku infrastrukturu. Gradovi imaju manju prosječnu stopu nezaposlenosti što je također bilo za očekivati s obzirom na to da je u gradovima veća mogućnost zaposlenja. Posljednje, gradovi imaju veći prosječni dohodak po stanovniku što je posljedica visokoplaćenih poslova koji se najčešće nalaze u gradovima.

U ovom poglavlju navedene su varijable koje će biti korištene u radu, njihov izvor i statistička obilježja radi boljeg razumijevanja. Posebno, za svaku zavisnu varijablu prikazana je aritmetička sredina u odnosu na pojedinu razinu transparentnosti te se na temelju toga moglo doći do pretpostavke koje hipoteze će biti odbačene. Dodatno, ovo poglavlje je važno je kako bi se mogli odrediti modeli koji će se koristiti pri testiranju hipoteza.

Poglavlje 3

Metodologija

Za analizu utjecaja jedne ili više nezavisnih varijabli na zavisnu varijablu koriste se regresijske metode. Linearna regresija najosnovnija je regresijska metoda. Transparentnost je kategorijalna varijabla s vrijednostima od 0 do 5, dok su prihod, rashod, suficit i javnih dug pozitivne neprekidne varijable. Stoga će kod testiranja hipoteza biti korištene linearna regresija te dodatno višestruka linearna regresija čije će osnovne značajnosti i pretpostavke biti opisane u ovom poglavlju. Ovo poglavlje napisano je na temelju [11] i [17].

3.1 Jednostavna linearna regresija

Neka je x nezavisna varijabla te y zavisna varijabla. Njihov odnos se linearnom regresijom opisuje kao:

$$y = \beta_0 + \beta_1 x + \epsilon$$

gdje su β_1 i β_2 parametri modela, a ϵ slučajna pogreška. Za n opažanja, model linearne regresije zapisuje se kao:

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \; i = 1, ..., n$$

Parametar β_0 predstavlja konstantni član regresijske jednadžbe te kako predstavlja presjek pravca s y-osi, nazivamo ga odsječak. Parametar β_1 predstavlja nagib regresijske jednadžbe jer se za svaku jedinicu porasta u x, y poveća za β_1. Slučajne pogreške ϵ_i predstavljaju razliku između promatranih y_i-ieva i njihovih predviđenih vrijednosti $\hat{y}_i$.

3.2 Višestruka linearna regresija

U slučaju da utjecaj nezavisne varijable na zavisnu nije statistički značajan, nekada je potrebno uključiti kontrolne varijable u model. Tada koristimo višestruku linearnu regresiju koja se temelji na jednostavnoj linearnoj regresiji.

Neka su $x_1, ..., x_k$ nezavisne varijable te y zavisna varijabla. Model s k nezavisne varijable zapisuje se kao:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_k x_k + \epsilon,$$

gdje su $\beta_0, \beta_1, ..., \beta_k$ parametri modela , a ϵ slučajna pogreška.
Za n opažanja, imamo:

$$\begin{aligned} y_1 &= \beta_0 + \beta_1 x_{11} + \beta_1 x_{12} + ... + \beta_k x_{1k} + \epsilon_1, \\ &\;\;... \\ y_n &= \beta_0 + \beta_1 x_{n1} + \beta_n x_{n2} + ... + \beta_k x_{nk} + \epsilon_k, \end{aligned}$$

Odnosno, kada to zapišemo u matričnoj formi:

$$Y = X\beta + \epsilon,$$

gdje su: $Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{pmatrix}$, $X = \begin{pmatrix} 1 & x_{11} & x_{12} & x_{13} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & x_{23} & \dots & x_{2k} \\ 1 & x_{31} & x_{32} & x_{33} & \dots & x_{3k} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & x_{n3} & \dots & x_{nk} \end{pmatrix}$, $\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix}$ te $\epsilon = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \vdots \\ \epsilon_n \end{pmatrix}$

U matričnom zapisu Y nazivamo vektorom odgovora, X matricom dizajna, β vektorom parametara, a ϵ vektorom greške.

3.3 Procjena parametara

U ovom poglavlju definirat će se procjenitelji za $\hat{\beta_0}$ i $\hat{\beta_1}$ jednostavne linearne regresije, odnosno procjenitelji $\hat{\beta_i}$ za *i=1,...,k* višestruke linearne regresije. Oni će se procijeniti metodom najmanjih kvadrata (OLS, tj. *ordinary least square*).

3.3.1 Procjena parametra jednostavne linearne regresije

Kako bismo dobili procjene $\hat{\beta_0}$ i $\hat{\beta_1}$ parametara β_0 i β_0 koristimo metodu najmanjih kvadrata. Neka su (x_i, y_i) za $i = 1, ..., n$ nezavisna opažanja. Slijedi da je procijenjena vrijednost $\hat{y_i}$ jednaka:

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i,$$

a i-ta pogreška jednaka:

$$\epsilon_i = y_i - \hat{y}_i.$$

Metoda najmanjih kvadrata procjenjuje parametre β_0 i β_1 tako da minimizira sumu kvadrata grešaka:

$$\begin{aligned} L(\beta_0, \beta_1) &= \sum_{i=1}^{n} \epsilon_i^2 \\ &= \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \\ &= \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) \end{aligned}$$

Sada slijedi da je parcijalna derivacija obzirom na β_0 i β_1 jednaka:

$$\frac{\partial L}{\partial \hat{\beta}_0}(\beta_0, \beta_1) = -2 \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) \tag{3.3.1}$$

$$\frac{\partial L}{\partial \hat{\beta}_1}(\beta_0, \beta_1) = -2 \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) x_i. \tag{3.3.2}$$

Izjednačavanjem (3.3.1) s nula dobivamo:

$$\sum_{i=1}^{n} y_i - n\hat{\beta}_0 - \hat{\beta}_1 \sum_{i=1}^{n} x_i = 0,$$

a (3.3.2) s nula:

$$\sum_{i=1}^{n} y_i x_i - n\hat{\beta}_0 x_i - \hat{\beta}_1 \sum_{i=1}^{n} x_i^2 = 0$$

Ubacivanjem

$$\overline{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \quad \text{i} \quad \overline{y} = \frac{1}{n} \sum_{i=1}^{n} y_i$$

u (3.3.1) dobivamo:

$$\hat{\beta}_0 = \overline{y} - \hat{\beta}_1 \overline{x} \tag{3.3.3}$$

dok ubacivanjem u (3.3.2) dobivamo:

$$\sum_{i=1}^{n} y_i x_i - n\hat{\beta}_0 \overline{x} - \hat{\beta}_1 \sum_{i=1}^{n} x_i^2 = 0. \tag{3.3.4}$$

Uvrštavanjem (3.3.3) u (3.3.4) slijedi

$$\sum_{i=1}^{n} y_i x_i - n\overline{x}(\overline{y} - \widehat{\beta_1 \overline{x}}) - \hat{\beta}_1 \sum_{i=1}^{n} x_i^2 = 0.$$

odnosno,

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} x_i y_i - n\overline{xy}}{\sum_{i=1}^{n} x_i^2 - n\overline{x}^2}. \tag{3.3.5}$$

Sada je još potrebno dokazati da su procjenitelji $\hat{\beta}_0$ i $\hat{\beta}_1$ određeni s (3.3.3) i (3.3.4) zaista minimumi funkcije $L(\hat{\beta}_0, \hat{\beta}_1)$.
Znamo da za minimume funkcije tada treba biti zadovoljeno:

$$\frac{\partial^2 L}{\partial^2 \hat{\beta}_0}(\hat{\beta}_0, \hat{\beta}_1) > 0$$

i

$$\frac{\partial^2 L}{\partial^2 \hat{\beta}_0}(\hat{\beta}_0, \hat{\beta}_1) \frac{\partial^2 L}{\partial^2 \hat{\beta}_1}(\hat{\beta}_0, \hat{\beta}_1) - \Big(\frac{\partial^2 L}{\partial \hat{\beta}_0 \partial \hat{\beta}_1}(\hat{\beta}_0, \hat{\beta}_1) \Big)^2 > 0$$

Kako je

$$\frac{\partial^2 L}{\partial^2 \hat{\beta}_0}(\hat{\beta}_0, \hat{\beta}_1) = \frac{\partial}{\partial \hat{\beta}_0} \Big(-2 \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) \Big) = 2n > 0$$

$$\frac{\partial^2 L}{\partial^2 \hat{\beta}_1}(\hat{\beta}_0, \hat{\beta}_1) = \frac{\partial}{\partial \hat{\beta}_1} \Big(-2 \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) x_i \Big) = 2 \sum_{i=1}^{n} x_i^2$$

$$\frac{\partial^2 L}{\partial\hat{\beta}_0\partial\hat{\beta}_1}(\hat{\beta}_0,\hat{\beta}_1) = \frac{\partial}{\partial\hat{\beta}_1}(-2\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 2\sum_{i=1}^{n} x_i$$

ubacivanjem u gornje formule slijedi da su $\hat{\beta}_0$ i $\hat{\beta}_1$ minimumi funkcije *SSE*.

Propozicija 3.3.1. *Procjenitelj $\hat{\beta}_1$ može se zapisati i kao*

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^{n}(x_i - \overline{x})^2} \tag{3.3.6}$$

Dokaz. Budući da vrijedi $\sum_{i=1}^{n}(y_i - \overline{y}) = 0$, slijedi

$$\begin{aligned}\sum_{i=1}^{n}(y_i - \overline{y})(x_i - \overline{x}) &= \sum_{i=1}^{n}(y_i x_i - \overline{y}x_i) - \overline{x}\sum_{i=1}^{n}(y_i - \overline{y})\\ &= \sum_{i=1}^{n} y_i x_i - \overline{y}\sum_{i=1}^{n} x_i\\ &= \sum_{i=1}^{n} x_i y_i - n\overline{y}\overline{x}\end{aligned}$$

te kako je

$$\begin{aligned}\sum_{i=1}^{n}(x_i - \overline{x})^2 &= \sum_{i=1}^{n} x_i^2 - 2\overline{x}\sum_{i=1}^{n} x_i + n\overline{x}^2\\ &= \sum_{i=1}^{n} x_i^2 - 2n\overline{x}^2 + n\overline{x}^2\\ &= \sum_{i=1}^{n} x_i - n\overline{x}^2.\end{aligned}$$

Ubacivanjem u (3.3.5) slijedi tvrdnja. □

3.3.2 Procjena parametara višestruke linearne regresije

Neka su $(x_{i1}, x_{i2}, ..., x_{ik}, y_i)$ za i = *1,...,n* nezavisna opažanja te b procjenitelj od β. Tada je procjenitelj $\hat{Y}$ jednak:

$$\hat{Y} = Xb,$$

te je

$$\epsilon = Y - \hat{Y}$$

Definicija 3.3.2. *Kažemo da je kvadratna matrica A simetrična ako vrijedi $A' = A$.*

Minimiziramo sumu kvadrata reziduala:

$$\begin{aligned} S &= (Y - \hat{Y})'(Y - \hat{Y}) \\ &= (Y - X\beta)'(Y - X\beta) \\ &= YY' - \beta' X' Y - Y' X\beta + \beta' X' X\beta \\ &= YY' - 2\beta'(X'Y) + \beta' X' X\beta \end{aligned}$$

gdje zadnja jednakost slijedi iz činjenice da je $Y'X\beta$ skalar.
Sada, koristeći diferencijaciju matrica dobivamo:

$$\frac{\partial S}{\partial \beta} = -2X'Y + 2X'X\beta, \tag{3.3.7}$$

gdje smo koristili da je $X'X$ simetrična matrica.

Definicija 3.3.3. *Kaže se da je matrica $A \in M_n(F)$ regularna ako postoji matrica $B \in M_n(F)$ takva da vrijedi $AB = BA = I$. U tom slučaju se matrica B zove multiplikativni inverz ili inverzna matrica od A i označava se s A^{-1}.*

Teorem 3.3.4. *Neka je $n \geqslant k$ i X $n \times k$ matrica punog stupčanog ranga, tada je X'X $k \times k$ pozitivno definitna matrica.*

Dokaz. Trebamo pokazati da je $y'X'Xy > 0$ za $y \neq 0$. Primijetimo da je $y'X'Xy = (Xy)'(Xy)$ što je pozitivno osim ako je $Xy = 0$. Budući da je X punog ranga, slijedi $Xy \neq 0$ za sve $y \neq 0$. Od tuda slijedi tvrdnja. □

Kako iz Teorema 3.3.4 slijedi da je $X'X$ regularna, izjednačavanjem izraza (3.3.7) s nula dobivamo da je kandidat za procjenitelja od β, u oznaci b jednak:

$$b = (X'X)^{-1}X'Y. \tag{3.3.8}$$

Sada još moramo dokazati da S postiže minimum u b. Zapišimo S na sljedeći način:

$$\begin{aligned}
S(\beta) &= (Y - X\beta)'(Y - X\beta) \\
&= (Y - Xb + Xb - X\beta)'(Y - Xb + Xb - X\beta) \\
&= (Y - Xb)'(Y - Xb) + 2(Y - Xb)'(Xb - X\beta) + (Xb - X\beta)'(Xb - X\beta) \\
&= (Y - Xb)'(Y - Xb) + 2(X(b - \beta))'(Y - Xb) + (X(b - \beta))'(X(b - \beta)) \\
&= (Y - Xb)'(Y - Xb) + 2(b - \beta)'X'(Y - Xb) + (b - \beta)'X'X(b - \beta) \\
&= (Y - Xb)'(Y - Xb) + 2(b - \beta)'(X'Y - X'Xb) + (b - \beta)'X'X(b - \beta) \\
&= (Y - Xb)'(Y - Xb) + v'v
\end{aligned}$$

gdje smo u zadnjoj jednakosti iskoristili (3.3.8), a $X(b - \beta)$ zapisali kao v.
Kako je $v'v \geq 0$ uz jednakost akko je $v = X(b - \beta) = 0$ što vrijedi akko $\beta = b$ budući da je X punog ranga. Dakle, $S(\beta) \geq S(b)$ uz jednakost akko je $\beta = b$.

Definicija 3.3.5. *Kažemo da je $\hat{\beta}_j$ nepristrani procjenitelj za β_j ako vrijedi $\mathbb{E}[\hat{\beta}_j] = \beta_j$.*

3.4 Gauss-Markovljevi uvjeti

U ovom poglavlju navest će se uvjeti potrebni kako bi se nepoznati parametri β_i, za *i=1,...,k* mogli procijeniti metodom najmanjih kvadrata. Pretpostavke koje moraju biti zadovoljene, poznate kao Gauss-Markovljevi uvjeti su sljedeće:

1. $\mathbb{E}[\epsilon_i] = 0$, za sve $i = 1, ...n$
2. $Var(\epsilon_i) = \sigma^2 > 0$, za sve $i = 1, ..., n$
3. $\mathbb{E}[\epsilon_i \epsilon_j] = 0$, za sve $i, j = 1, ..., n$ takve da je $i \neq j$

odnosno, u matričnoj formi:

1. $\mathbb{E}[\epsilon] = o$
2. $Cov(\epsilon) = \sigma^2 I$,

gdje je o nul-vektor. Primijetimo da prema gore navedenim uvjetima slijedi:

$$\mathbb{E}[Y] = X\beta \tag{3.4.1}$$

te

$$Cov(Y) = \mathbb{E}[(Y - X\beta)(Y - X\beta)'] = \mathbb{E}[\epsilon\epsilon'] = \sigma^2 I \tag{3.4.2}$$

Teorem 3.4.1. *Pod Gauss-Markovljevim uvjetima vrijedi da je procjenitelj b dobiven metodom najmanjih kvadrata nepristran procjenitelj od β, odnosno $\mathbb{E}[b] = \beta$. Također, vrijedi $Cov(b) = \sigma^2(X'X)^{-1}$.*

Dokaz.

$$\begin{aligned}\mathbb{E}[b] &= \mathbb{E}[(X'X)^{-1}X'Y] \\ &= (X'X)^{-1}X'\mathbb{E}[Y] \\ &= (X'X)^{-1}X'X\beta \\ &= \beta\end{aligned}$$

gdje je u trećoj jednakosti korišten (3.4.1).

$$\begin{aligned}Cov(b) &= Cov((X'X)^{-1}X'Y) \\ &= (X'X)^{-1}X'Cov(Y)((X'X)^{-1}X')' \\ &= (X'X)^{-1}X'Cov(Y)X((X'X)^{-1}) \\ &= (X'X)^{-1}X'(\sigma^2 I)X((X'X)^{-1}) \\ &= \sigma^2(X'X)^{-1}X'X(X'X)^{-1} \\ &= \sigma^2(X'X)^{-1}\end{aligned}$$

gdje je u četvrtoj jednakosti korišteno (3.4.2). □

Kako bismo pronašli procjenitelja za σ^2 treba nam sljedeći teorem.

Teorem 3.4.2. *Neka je y n×1 vektor i A n×n simetrična matrica, tada*

$$\mathbb{E}(y'Ay) = tr(ACov(y)) + \mathbb{E}(y)'A\mathbb{E}(y)$$

Dokaz. Dokaz se nalazi u [17] □

Definicija 3.4.3. *Procjenitelj $\hat{\beta}_j$ konzistentan je procjenitelj za β_j ako povećanjem uzorka konvergira po vjerojatnosti prema stvarnoj vrijednosti, to jest ako vrijedi*

$$\lim_{N\to\infty} \mathbb{P}(|\hat{\beta}_j - \beta_j| < \epsilon) = 1, \forall \epsilon > 0$$

Teorem 3.4.4. *Pod Gauss-Markovljevim uvjetima,*

$$s^2 = \frac{S(b)}{n-k-1} \tag{3.4.3}$$

nepristran je i konzistentan procjenitelj za σ^2.

Dokaz.

$$\begin{aligned} S(b) &= (Y - Xb)'(Y - Xb) \\ &= Y'Y - 2b'X'Y + b'X'Xb \\ &= Y'Y - 2bX'Y + bX'Y \\ &= Y'Y - b'X'Y \\ &= Y'Y - Y'X(X'X)^{-1}X'Y \\ &= Y'Y - Y'HY \\ &= Y'(I - H)Y \end{aligned}$$

gdje je $H = X(X'X)^{-1}X'$. Pokažimo da je $(H - I)$ simetrična matrica:

$$\begin{aligned} (H - I)' &= [X(X'X)^{-1}X']' - I' \\ &= X[(X'X)']^{-1}X - I \\ &= X(X'X)^{-1}X - I \\ &= H - I \end{aligned}$$

Sada, koristeći Teorem 3.4.2, (3.4.1) i (3.4,2) slijedi

$$\begin{aligned} \mathbb{E}[S(b)] &= tr((I_n - H)(\sigma^2 I)) + (Xb)'(I - H)(Xb) \\ &= \sigma^2 tr(I_n - H) + b'X'(I - H)Xb \\ &= \sigma^2 tr(I_n - H) + b'(X'X - X'X(X'X)^{-1}X'X)b \\ &= \sigma^2 tr(I_n - H) + b'(X'X - X'X)b \\ &= \sigma^2 tr(I_n - H) \\ &= \sigma^2 (tr(I_n) - tr(H)) \end{aligned}$$

Koristeći činjenicu da je $tr(AB) = tr(BA)$ slijedi

$$\begin{aligned} tr(H) &= tr(X(X'X)^{-1}X') \\ &= tr(X'X(X'X)^{-1}) \\ &= tr(I_{k+1}) \\ &= k + 1 \end{aligned}$$

Dakle,

$$\begin{aligned} \mathbb{E}[S(b)] &= \sigma^2(n - (k + 1)) \\ &= \sigma^2(n - k - 1) \end{aligned}$$

Prema tome, dobivamo da je s^2 nepristran procjenitelj za σ^2.
Također, kada $n \to \infty$ tada $s^2 \to \sigma^2$. Dakle, s^2 je konzistentan procjenitelj za σ^2. $\square$

U sljedećem teoremu spominje se najbolji nepristran linearni procjenitelj (BLUE, tj. best linear unbiased estimator) , gdje najbolji znači da ima najmanju varijancu, a linearan da se može prikazati kao linearna funkcija $y_1, ..., y_n$, odnosno kao $c'Y$.

Teorem 3.4.5. *(Gauss-Markovljev teorem) Neka je $b = (XX')^{-1}Y$ te $Y = X\beta + \epsilon$. Tada je uz Gauss-Markovljeve uvjete, procjenitelj $a'b$ BLUE od $a'\beta$*

Dokaz. Neka je $c'Y$ nepristran procjenitelj za $a'\beta$. Tada imamo $\mathbb{E}[c'Y] = c'\mathbb{E}[Y] = c'X\beta = a'\beta$ što zahtjeva

$$c'X = a'. \tag{3.4.4}$$

Uz (3.4.2) slijedi da je

$$Var(c'Y) = c'Cov(Y)c = \sigma^2 c'c$$

odnosno,

$$\begin{aligned} Var(c'Y) &= \sigma^2[c - X(X'X)^{-1}a + X(X'X)^{-1}a]'[c - X(X'X)^{-1}a + X(X'X)^{-1}a] \\ &= \sigma^2\Big\{[c - X(X'X)^{-1}a]'[c - X(X'X)^{-1}a] + 2[c - X(X'X)^{-1}a]'[X(X'X)^{-1}a] \\ &\quad + [X(X'X)^{-1}a]'[X(X'X)^{-1}a]\Big\} \end{aligned}$$

sada uz (3.3.4) imamo

$$Var(c'Y) = \sigma^2 \Big\{ [c - X(X'X)^{-1}a]'[c - X(X'X)^{-1}a] + a'(X'X)^{-1}a \Big\}$$

što postiže minimum za $c = XX(X'X)^{-1}a$. Odnosno, za takav c slijedi da je $c'Y$ jednak $a'b$. □

3.5 Validacija modela

Kako bismo utvrdili je li model prikladan, potrebno je napraviti njegovu validaciju. To uključuje provjeru koliko predviđene vrijednosti modela odstupaju od stvarnih, kolika je statistička značajnost parametara te provjeru o normalnoj distribuciji i nezavisnosti slučajnih pogrešaka.

3.5.1 Koeficijent determinacije

Koeficijent determinacije R^2 mjera je koja pokazuje koliki postotak varijance zavisne varijable se objašnjava prediktorskim varijablama. Koeficijent determinacije poprima vrijednosti iz intervala [0,1] te je zadan formulom:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{\sum_{i=1}^{n} (\overline{y_i} - y_i)^2}$$

Za R^2=0 vrijedi da predviđene vrijednosti $\hat{y}_i$ odgovaraju sredini uzorka $\overline{y_i}$. S druge strane, za R^2=1 slijedi da su predviđene vrijednosti $\hat{y}_i$ i stvarne vrijednosti y_i jednake. Dakle, što je R^2 bliži 1, to je prilagodba modela bolja.

Uvođenjem dodatnih prediktorskih varijabli, R^2 se neće smanjiti, neovisno o tome da li dodane varijable lošije opisuju zavisnu varijablu. Stoga se u analizama se često koristi i prilagođeni R^2, u oznaci R_a^2 koji je dan s:

$$R_a^2 = 1 - \frac{\frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{n-k-1}}{\frac{\sum_{i=1}^{n} (\overline{y_i} - y_i)^2}{n-1}} = 1 - \frac{n-1}{n-k-1}(1 - R^2)$$

te koji se prilagođava broju prediktorskih varijabli. Prilagođeni R_a^2 se smanjuje dodavanjem novih varijabli, za razliku od običnog R^2.

3.5.2 F - test

F - testom testiramo postoji li jedan barem jedan regresijski parametar β_i različit od nule, odnosno hipoteze su sljedeće:

$$H_0 : \beta_0 = \beta_1 = ... = \beta_k = 0$$
$$H_1 : \exists i \in \{1, ..., n\}\ t.d.\ \beta_i \neq 0$$

Ukoliko bi se nul-hipoteza pokazala istinitom, to bi značilo da je model zapravo jednak $Y = \beta_0 + \epsilon$. Definiramo sumu kvadrata reziduala modela $Y = \beta_0 + \epsilon$, RRS_1, kao:

$$RRS_1 = \sum_{i=1}^{n} (\beta_0 - y_i)^2,$$

te sumu kvadrata reziduala dobivenog modela, RRS_2, kao:

$$RRS_2 = \sum_{i=1}^{n} (\hat{y}_i - y_i)^2.$$

Sada je testna statistika dana s:

$$F = \frac{n-k-1}{k} \frac{RRS_1 - RRS_2}{RRS_1}$$

te ima distribuciju jednaku:

$$F \sim F(k, n-k-1)$$

3.5.3 t - test

Za testiranje značajnosti parametara modela koristi se sljedeće hipoteze:

$$H_0 : \beta_j = b_j$$
$$H_1 : \beta_j \neq b_j$$

za $j \in 1, ..., k$.
Ovakav test nazivamo dvostrani t - test budući da prema H_1, β_j može biti manji ili veći od

b_j.
Testna statistika definirana je s:

$$T = \frac{\hat{\beta}_j - b_j}{\hat{\sigma}\sqrt{(X'X)^{-1}_{jj}}} \sim t(n-k-1)$$

Dodatno, $(1-\alpha)\%$ pouzdani interval za β_j dan je s:

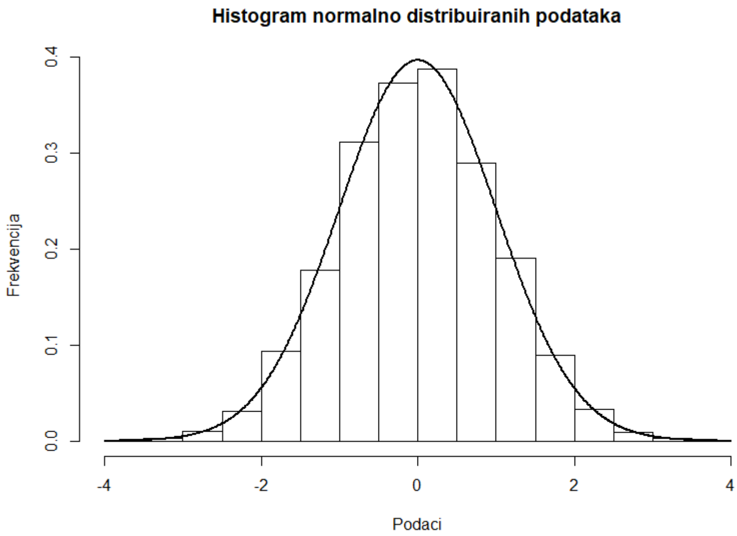
$$\left\langle \hat{\beta}_j - t_{1-\frac{\alpha}{2},n-k-1}\hat{\sigma}\sqrt{(X'X)^{-1}_{jj}}, \hat{\beta}_j + t_{1-\frac{\alpha}{2},n-k-1}\hat{\sigma}\sqrt{(X'X)^{-1}_{jj}} \right\rangle$$

gdje $t_{1-\frac{\alpha}{2},n-k-1}$ predstavlja $(1-\frac{\alpha}{2})$ kvantil Studentove t-distribucije s (n-k-1) stupanj slobode. Posebno, za $b_j = 0$ testira se ima li varijabla X_j utjecaj na Y. U tom slučaju, ako odbacimo nul-hipotezu, kažemo da X_j statistički značajno utječe na Y.

3.5.4 Testiranje normalnosti pogrešaka

Prije korištenja metode najmanjih kvadrata, potrebno je provjeriti vrijede li uvjeti definirani u Potpoglavlju 3.4, odnosno $\epsilon_i \sim N(0, \sigma^2)$ t.d. ϵ_i, ϵ_j nezavisne za svaki $i \neq j$. Normalnost distribucije možemo testirati grafički te koristeći Lillieforsovu inačicu Kolmogorov-Smirnovljeva testa.

Grafičko testiranje daje lako čitljiv pregled distribucije. Crtanje histograma i normalnog vjerojatnosnog grafa najčešći su grafički testovi normalnosti uzorka.
Histogram zvonolikog oblika sa srednjom vrijednošću oko 0, pokazatelj je da bi uzorak mogao biti iz normalne distribucije s očekivanjem 0. Na slici 3.1 prikazan jedan takav histogram.



Slika 3.1: Histogram s funkcijom gustoće normalne razdiobe

Izvor: izrada autora

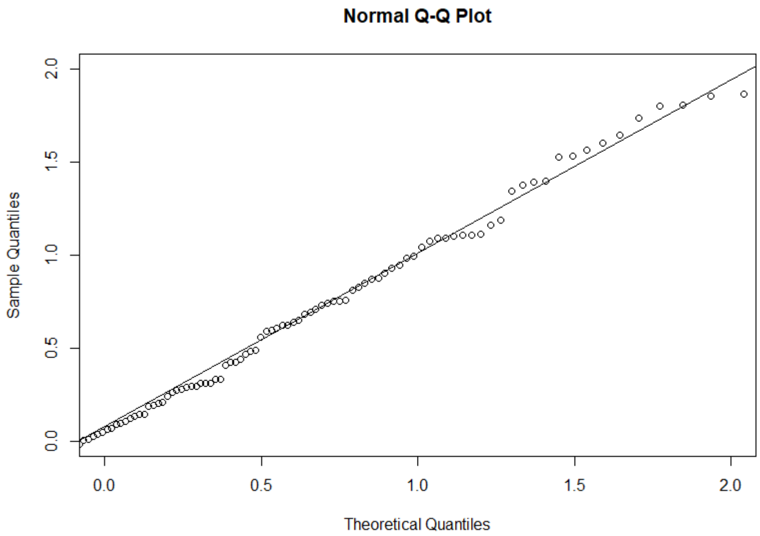
Normalni vjerojatnosni graf ili $Q - Q$ *plot* (*quantile* $-$ *quantile*) crta se tako da se sortirane slučajne pogreške crtaju uz kvantile normalne distribucije. Odnosno, prikazujemo

$$(x_i, \epsilon^i), \ za\ i = 1, ..., n$$

gdje su

$$x_i = \phi^{-1}\left(\frac{i - \frac{3}{8}}{n + \frac{1}{4}}\right).$$

kvantili standardne normalne distribucije budući da je ϕ^{-1} inverzna funkcija jedinične normalne distribucije, a ϵ^i sortirane slučajne pogreške.
Ako točke prate pravac koji siječe y-os u 0, tada možemo zaključiti da su slučajne pogreške normalno distribuirane s očekivanjem 0. Primjer normalnog vjerojatnosnog grafa koji prikazuje upravo to dan je na slici 3.2.



Slika 3.2: Normalni vjerojatnosni graf

Izvor: izrada autora

Normalnost slučajnih pogrešaka osim grafički testiramo i Lilliforsovom inačicom Kolmogorov-Smirnovljeva testa.
Hipoteze koje se testiraju su sljedeće:

$$H_0 : F = F_0$$
$$H_1 : F \neq F_0$$

gdje je F distribucija uzorka $X_1, ..., X_n$, a F_0 standardna normalna distribucija.
Testna statistika dana je s:

$$D_n = \underset{x \in \mathbb{R}}{sup} |\hat{F}_n - F_0(x)| = \underset{1 \leq i \leq n}{max} \{ max\{ |\frac{i-1}{n} - F_0(x_i)|, |\frac{i}{n} - F_0(x_i)| \} \}$$

Nas zanima pripadaju li slučajne pogreške normalnoj distribuciji, stoga će nam F biti distribucija slučajnih pogrešaka normiranih na sljedeći način:

$$\epsilon_i' = \frac{\epsilon_i}{\hat{\sigma}}$$

gdje je $\hat{\sigma}$ procijenjena standardna devijacija slučajnih pogrešaka. Tada nam empirijska funkcija distribucije normiranih slučajnih pogrešaka definirana s:

$$\hat{F}_n = \frac{1}{n} \sum_{i=1}^{n} 1_{\{\epsilon_i' \leq x\}}$$

Vrijednost testne statistike D_n najveća je apsolutna vrijednost odstupanja empirijske funkcije distribucije slučajnih pogrešaka od standardne normalne distribucije.
Kritično područje je:

$$[d_\alpha(n), \infty >$$

gdje je α razina značajnosti, a $d_\alpha(n)$ vrijednost iščitana iz Lillieforsove tablice. Ako vrijednost testne statistike D_n upada u kritično područje, tada odbacujemo nultu hipotezu.

U ovom poglavlju opisana je metodologija koja će biti korištena za testiranje hipoteza. Budući da je proračunska transparentnost diskretna varijabla, a proračunski rashod, prihod, suficit i javni dug po stanovniku neprekidne varijable, za testiranje hipoteza koristit će se jednostavna linearna regresija. Uvođenjem kontrolnih varijabli hipoteze će se testirati višestrukom linearnom regresijom. Ovo poglavlje sadrži osnovne informacije o spomenutim modelima, njihovim procjeniteljima, kao i pretpostavkama koje moraju biti zadovoljene da bi se oni mogli koristiti. Nakon opisa metodologije slijedi testiranje hipoteza.

Poglavlje 4

Testiranje hipoteza

U ovom poglavlju provodi se empirijska analiza na podacima koristeći metodologiju iz Poglavlja 3. Proračunska transparentnost nezavisna je varijabla označena s x, proračunski rashod po stanovniku, proračunski prihod po stanovniku, proračunski suficit po stanovniku i javni dug po stanovniku zavisne su varijable označene s y. Kontrolne varijable označene su kao u Poglavlju 2.

4.1 Testiranje Hipoteze 1

Koristeći linearnu regresiju testirat će se prva hipoteza, odnosno doprinosi li veća transparentnost povećanju proračunskog rashoda. Za analizu podataka koristi se naredba *lm* u R Studio-u.

4.1.1 Testiranje Hipoteze 1 na gradovima

Kako bi se zadovoljile pretpostavke linearne regresije, proračunski rashod kod gradova transformiran je koristeći logaritam. Primjenom modela jednostavne linearne regresije na transformiranim podacima dobiven je sljedeći model:

$$ln(y) = 8{,}90433 - 0{,}04815x$$

odnosno,

$$y = e^{8{,}90433 - 0{,}04815x}$$

U Tablici 4.1 dani su rezultati linearne regresije između proračunske transparentnosti i rashoda po stanovniku u gradovima. Prema koeficijentu uz transparentnost slijedi da povećanje transparentnosti dovodi do smanjenja rashoda, suprotno istraživanjima navedenim u Poglavlju 1, no prema p-vrijednosti koeficijenta uz transparentnost slijedi da on nije

statistički značajan.
Koeficijent determinacije $R^2 = 0{,}0116$, odnosno samo 1,2% rashoda po stanovniku gradova može se objasniti transparentnošću. Stoga možemo zaključiti da ne postoji veza između proračunske transparentnosti i rashoda po stanovniku u gradovima.

Koeficijent	**Procjena koeficijenta**	**Standardna devijacija**	**t - test statistika**	**p - vrijednost**
β_0	8,90433	0,18152	49,055	<2e-16
β_1	-0,04815	0,03960	-1,216	0,226

Tablica 4.1: Rezultati linearne regresije između proračunske transparentnosti i rashoda po stanovniku gradova

4.1.2 Testiranje Hipoteze 1 na općinama

Kako bi se zadovoljile pretpostavke linearne regresije, proračunski rashod po stanovniku kod općina transformiran je koristeći logaritam. Primjenom modela jednostavne linearne regresije na transformiranim podacima dobiven je sljedeći model:

$$ln(y) = 8{,}464284 - 0{,}001137x$$

odnosno,

$$y = e^{8{,}464284 - 0{,}001137x}$$

U tablici 4.2 dani su rezultati linearne regresije između proračunske transparentnosti i rashoda po stanovniku gradova. Prema koeficijentu uz transparentnost općina slijedi da povećanje transparentnosti dovodi do smanjenja rashoda, suprotno istraživanjima navedenim u Poglavlju 1, no prema p-vrijednosti koeficijenta uz transparentnost slijedi da on nije statistički značajan.
Koeficijent determinacije $R^2 = 0{,}0005$, odnosno samo 0,05% rashoda po stanovniku u gradova može se objasniti transparentnošću. Stoga, možemo zaključiti da ne postoji veza između proračunske transparentnosti i rashoda po stanovniku u gradovima.

Koeficijent	**Procjena koeficijenta**	**Standardna devijacija**	**t - test statistika**	**p - vrijednost**
β_0	8,464284	0,069149	122,407	<2e-16
β_1	-0,001137	0,016492	-0,069	0,945

Tablica 4.2: Rezultati linearne regresije između proračunske transparentnosti i rashoda po stanovniku općina

4.1.3 Testiranje Hipoteze 1 s kontrolnim varijablama na gradovima

Kako bi se zadovoljile pretpostavke linearne regresije, proračunski rashod kod gradova transformiran je koristeći logaritam. Primjenom modela višestruke linearne regresije na transformiranim podacima dobiven je sljedeći model:

$$ln(y) = 7{,}940 - 0{,}012x - 0{,}464I + 1{,}141N + 0{,}00003D$$

odnosno,

$$y = e^{7{,}940-0{,}012x-0{,}464I+1{,}141N+0{,}00003D}$$

U Tablici 4.3 dani su rezultati višestruke linearne regresije između rashoda gradova i gore navedenih nezavisnih varijabli. Prema koeficijentu uz transparentnost slijedi da povećanje transparentnosti dovodi do smanjenja rashoda, suprotno istraživanjima navedenim u Poglavlju 1, no prema p-vrijednosti koeficijenta uz transparentnost slijedi da on nije statistički značajan. Statistički su značajni koeficijenti uz stopu nezaposlenosti i dohodak te njihovo povećanje dovodi do povećanja rashoda.
Prilagođeni koeficijent determinacije $R^2 = 0{,}0649$, odnosno 6,5% rashoda po stanovniku u gradova može se objasniti nezavisnim varijablama. F - statistika je jednaka 10,63, a p - vrijednost modela 2,635e-10, stoga možemo zaključiti da je model statistički značajan.

Koeficijent	**Procjena koeficijenta**	**Standardna devijacija**	**t - test statistika**	**p - vrijednost**
β_0	7,940	0,1635	48,565	<2e-16
β_1	-0,012	0,01616	-0,742	0,458
β_2	-0,464	0,2957	-1,571	0,117
β_3	1,141	0,2664	4,284	2,17e-05
β_4	0,00003	0,000004	6,350	4,51e-10

Tablica 4.3: Rezultati višestruke linearne regresije za prvu hipotezu na gradovima

4.1.4 Testiranje Hipoteze 1 s kontrolnim varijablama na općinama

Kako bi se zadovoljile pretpostavke linearne regresije, proračunski rashod kod općina transformiran je koristeći logaritam. Primjenom modela višestruke linearne regresije na transformiranim podacima dobiven je sljedeći model:

$$ln(y) = 7{,}896 - 0{,}0076x - 0{,}6994I + 1{,}361N + 0{,}00003D$$

odnosno,

$$y = e^{7{,}896-0{,}0076x-0{,}6994I+1{,}361N+0{,}00003D}$$

U Tablici 4.4 dani su rezultati višestruke linearne regresije između rashoda općina i gore navedenih nezavisnih varijabli. Prema koeficijentu uz transparentnost slijedi da povećanje transparentnosti dovodi do smanjenja rashoda, suprotno istraživanjima navedenim u Poglavlju 1, no prema p-vrijednosti koeficijenta uz transparentnost slijedi da on nije statistički značajan. Koeficijenti uz stopu nezaposlenosti i dohodak statistički su značajni na svim razinama značajnosti, dok je koeficijent uz postotak korisnika interneta značajan na razini značajnosti do $\alpha = 5\%$. Povećanje stope nezaposlenosti i prosječnog dohotka dovodi do povećanja rashoda po stanovniku općina, dok povećanje postotka korisnika interneta dovodi do njegovog smanjenja.
Prilagođeni koeficijent determinacije R^2 = 0,07, odnosno 7% rashoda po stanovniku općina može se objasniti nezavisnim varijablama. F - statistika je jednaka 10,05, a p - vrijednost modela 5,105e-07, stoga možemo zaključiti da je model statistički značajan.

Koeficijent	**Procjena koeficijenta**	**Standardna devijacija**	**t - test statistika**	**p - vrijednost**
β_0	7,896	0,2051	38,492	<2e-16
β_1	-0,0076	0,01776	-0,428	0,6687
β_2	-0,6994	0,3447	-2,029	0,0431
β_3	1,361	0,2961	4,596	5,70e-06
β_4	0,00003	0,000005	5,963	5,23e-09

Tablica 4.4: Rezultati višestruke linearne regresije za prvu hipotezu na općinama

U ovom potpoglavlju testirala se hipoteza 1, odnosno testirao se utjecaj proračunske transparentnosti na proračunski rashod. Hipoteza 1 testirala se odvojeno na gradovima, odvojeno na općinama te dodatno uvođenjem kontrolnih varijabli. Svugdje se, suprotno istraživanjima, dobilo da povećanje transparentnosti dovodi do smanjenje rashoda, ali da koeficijent uz transparentnost nije statistički značajan. Kod gradova, 1,2% rashoda po stanovniku može se objasniti transparentnošću, dok se kod općina 0,05% rashoda po stanovniku može objasniti transparentnošću. Zaključak ovog potpoglavlja je da ne postoji veza između proračunske transparentnosti i rashoda po stanovniku hrvatskih općina i gradova.

4.2 Testiranje Hipoteze 2

Koristeći linearnu regresiju testirat će se druga hipoteza, odnosno doprinosi li veća transparentnost povećanju proračunskog prihoda.

4.2.1 Testiranje Hipoteze 2 na gradovima

Kako bi se zadovoljile pretpostavke linearne regresije, proračunski prihod kod gradova transformiran je koristeći logaritam. Primjenom modela jednostavne linearne regresije na transformiranim podacima dobiven je sljedeći model:

$$ln(y) = 8{,}74136 - 0{,}0063x$$

odnosno,

$$y = e^{8{,}74136 - 0{,}0063x}$$

U tablici 4.5 dani su rezultati linearne regresije između proračunske transparentnosti i prihoda po stanovniku gradova. Prema koeficijentu uz transparentnost slijedi da povećanje transparentnosti dovodi do smanjenja prihoda gradova, suprotno istraživanjima navedenim u Poglavlju 1, no prema p-vrijednosti koeficijenta uz transparentnost slijedi da on nije statistički značajan.
Koeficijent determinacije $R^2 = 0{,}0002$, odnosno samo 0,02% prihoda po stanovniku gradova može se objasniti transparentnošću. Stoga, možemo zaključiti da ne postoji veza između proračunske transparentnosti i prihoda po stanovniku u gradovima.

Koeficijent	**Procjena koeficijenta**	**Standardna devijacija**	**t - test statistika**	**p - vrijednost**
β_0	8,74136	0,17131	51,03	<2e-16
β_1	-0,00634	0,03737	-0,17	0,866

Tablica 4.5: Rezultati linearne regresije između proračunske transparentnosti i prihoda po stanovniku gradova

4.2.2 Testiranje Hipoteze 2 na općinama

Kako bi se zadovoljile pretpostavke linearne regresije, proračunski prihod kod općina transformiran je koristeći logaritam. Primjenom modela jednostavne linearne regresije na transformiranim podacima dobiven je sljedeći model:

$$ln(y) = 8{,}619103 - 0{,}009125x$$

odnosno,

$$y = e^{8{,}619103 - 0{,}009125x}$$

U tablici 4.6 dani su rezultati linearne regresije između proračunske transparentnosti i prihoda po stanovniku općina. Prema koeficijentu uz transparentnost slijedi da povećanje transparentnosti dovodi do smanjenja prihoda općina, suprotno istraživanjima navedenim u Poglavlju 1, no prema p-vrijednosti koeficijenta uz transparentnost slijedi da on nije

statistički značajan.
Koeficijent determinacije R^2 = 0,0006, odnosno samo 0.06% prihoda po stanovniku općina može se objasniti transparentnošću. Stoga, možemo zaključiti da ne postoji veza između proračunske transparentnosti i prihoda po stanovniku u općinama.

Koeficijent	**Procjena koeficijenta**	**Standardna devijacija**	**t - test statistika**	**p - vrijednost**
β_0	8,619103	0,073170	117,796	<2e-16
β_1	-0,009125	0,017483	-0,522	0,602

Tablica 4.6: Rezultati linearne regresije između proračunske transparentnosti i prihoda po stanovniku općina

4.2.3 Testiranje Hipoteze 2 s kontrolnim varijablama na gradovima

Kako bi se zadovoljile pretpostavke linearne regresije, proračunski prihod kod gradova transformiran je koristeći logaritam. Primjenom modela višestruke linearne regresije na transformiranim podacima dobiven je sljedeći model:

$$ln(y) = 8{,}083 - 0{,}0996x - 0{,}5865I + 1{,}056N + 0{,}00003D$$

odnosno,

$$y = e^{8{,}083-0{,}0996x-0{,}5865I+1{,}056N+0{,}00003D}$$

U Tablici 4.7 dani su rezultati višestruke linearne regresije između prihoda gradova i gore navedenih nezavisnih varijabli. Prema koeficijentu uz transparentnost slijedi da povećanje transparentnosti dovodi do smanjenja prihoda, suprotno istraživanjima navedenim u Poglavlju 1, no prema p-vrijednosti koeficijenta uz transparentnost slijedi da on nije statistički značajan. Koeficijenti uz stopu nezaposlenosti i dohodak statistički su značajni na svim razinama značajnosti, dok je koeficijent uz postotak korisnika interneta značajan na razini značajnosti do $\alpha = 5\%$. Povećanje stope nezaposlenosti i prosječnog dohotka dovodi do povećanja rashoda po stanovniku gradova, dok povećanje postotka korisnika interneta dovodi do njegovog smanjenja.
Prilagođeni koeficijent determinacije R^2 = 0,0722, odnosno 7,2% prihoda po stanovniku gradova može se objasniti nezavisnim varijablama. F - statistika je jednaka 10,72, a p - vrijednost modela 2,247e-08, stoga možemo zaključiti da je model statistički značajan.

Koeficijent	Procjena koeficijenta	Standardna devijacija	t - test statistika	p - vrijednost
β_0	8,083	0,1562	51,74	<2e-16
β_1	-0,0996	0,01545	-0,645	0,5193
β_2	-0,5865	0,2826	-2,076	0,0384
β_3	1,056	0,2546	4,150	3,85e-05
β_4	0,00003	0,000004	6,462	2,27e-10

Tablica 4.7: Rezultati višestruke linearne regresije za drugu hipotezu na gradovima

4.2.4 Testiranje Hipoteze 2 s kontrolnim varijablama na općinama

Kako bi se zadovoljile pretpostavke linearne regresije, proračunski prihoda općina transformiran je koristeći logaritam. Primjenom modela višestruke linearne regresije na transformiranim podacima dobiven je sljedeći model:

$$ln(y) = 8{,}055 - 0{,}0085x - 0{,}9045I + 1{,}361N + 0{,}00003D$$

odnosno,

$$y = e^{8{,}055-0{,}0085x-0{,}9045I+1{,}361N+0{,}00003D}$$

U Tablici 4.8 dani su rezultati višestruke linearne regresije između prihoda općina i gore navedenih nezavisnih varijabli. Prema koeficijentu uz transparentnost slijedi da povećanje transparentnosti dovodi do smanjenja prihoda, suprotno istraživanjima navedenim u Poglavlju 1, no prema p-vrijednosti koeficijenta uz transparentnost slijedi da on nije statistički značajan. Koeficijenti uz stopu nezaposlenosti i prosječni dohodak statistički su značajni na svim razinama značajnosti, dok je koeficijent uz postotak korisnika interneta značajan na razini značajnosti $\alpha = 1\%$. Povećanje stope nezaposlenosti i prosječnog dohotka dovodi do povećanja prihoda po stanovniku općina, dok povećanje postotka korisnika interneta dovodi do njegovog smanjenja.

Prilagođeni koeficijent determinacije $R^2 = 0{,}08$, odnosno 8% prihoda po stanovniku općina može se objasniti nezavisnim varijablama. F - statistika je jednaka 10,3, a p - vrijednost modela 5,734e-08, stoga možemo zaključiti da je model statistički značajan.

Koeficijent	Procjena koeficijenta	Standardna devijacija	t - test statistika	p - vrijednost
β_0	8,055	0,1969	40,916	<2e-16
β_1	-0,0085	0,01704	-0,496	0,62030
β_2	-0,9045	0,3308	-2,734	0,00651
β_3	1,322	0,2842	4,651	4,42e-06
β_4	0,00003	0,000005	6,347	5,66e-10

Tablica 4.8: Rezultati višestruke linearne regresije za drugu hipotezu na općinama

U ovom potpoglavlju testirala se hipoteza 2, koja je donesena na temelju prethodnih radova. Hipoteza 2 tvrdi da povećanje proračunske transparentnosti dovodi do povećanja proračunskog prihoda. Ona se jednostavnom linearnom regresijom testirala prvo za gradove, a onda za općine. Rezultati jednostavne linearne regresije suprotni su prethodnim istraživanjima, dobiveno je da povećanje transparentnosti dovodi do smanjenja prihoda po stanovniku. No, ta veza nije statistički značajna ni kod gradova, ni kod općina. Uvođenjem kontrolnih varijabli višestrukom linearnom regresijom dobiveni su jednaki rezultati. Zaključak ovog potpoglavlja je da ne postoji veza između proračunske transparentnosti i prihoda po stanovniku hrvatskih općina i gradova.

4.3 Testiranje Hipoteze 3

Koristeći linearnu regresiju testirat će se treća hipoteza, odnosno doprinosi li veća transparentnost ostvarenju proračunskog suficita.

4.3.1 Testiranje Hipoteze 3 na gradovima

Kako bi se zadovoljile pretpostavke linearne regresije, proračunski suficit kod gradova transformiran je koristeći logaritam. Primjenom modela jednostavne linearne regresije na transformiranim podacima dobiven je sljedeći model:

$$ln(y) = 5{,}92309 + 0{,}01428x$$

odnosno,

$$y = e^{5{,}92309+0{,}01428x}$$

U tablici 4.9 dani su rezultati linearne regresije između proračunske transparentnosti i suficita po stanovniku gradova. Prema koeficijentu uz transparentnost slijedi da povećanje transparentnosti dovodi do povećanja suficita gradova, kao i kod istraživanja navedenih u Poglavlju 1, no prema p-vrijednosti koeficijenta uz transparentnost slijedi da on nije statistički značajan.

Koeficijent determinacije R^2 = 8,43e-05, stoga, možemo zaključiti da ne postoji veza između proračunske transparentnosti i suficita po stanovniku gradova.

Koeficijent	Procjena koeficijenta	Standardna devijacija	t - test statistika	p - vrijednost
β_0	5,92309	0,82094	7,215	2,89e-10
β_1	0,01428	0,17493	0,082	0,935

Tablica 4.9: Rezultati linearne regresije između proračunske transparentnosti i suficita po stanovniku gradova

4.3.2 Testiranje Hipoteze 3 na općinama

Kako bi se zadovoljile pretpostavke linearne regresije, proračunski suficit kod općina transformiran je koristeći treći korijen. Primjenom modela jednostavne linearne regresije na transformiranim podacima dobiven je sljedeći model:

$$\sqrt[3]{y} = 8{,}73332 \text{ - } 0{,}07576x$$

odnosno,

$$y = (8,73332 - 0,07576x)^3$$

U tablici 4.10 dani su rezultati linearne regresije između proračunske transparentnosti i suficita po stanovniku općina. Prema koeficijentu uz transparentnost slijedi da povećanje transparentnosti dovodi do smanjenja suficita općina, suprotno istraživanjima navedenih u Poglavlju 1, no prema p-vrijednosti koeficijenta uz transparentnost slijedi da on nije statistički značajan.
Koeficijent determinacije R^2 = 0,001, odnosno samo 0,1% suficita po stanovniku općina može se objasniti transparentnošću. Stoga, možemo zaključiti da ne postoji veza između proračunske transparentnosti i suficitu po stanovniku općina.

Koeficijent	Procjena koeficijenta	Standardna devijacija	t - test statistika	p - vrijednost
β_0	8,73332	0,58978	14,808	<2e-16
β_1	-0,07576	0,14115	-0,537	0,592

Tablica 4.10: Rezultati linearne regresije između proračunske transparentnosti i suficita po stanovniku općina

4.3.3 Testiranje Hipoteze 3 s kontrolnim varijablama na gradovima

Kako bi se zadovoljile pretpostavke linearne regresije, proračunski suficit kod gradova transformiran je koristeći logaritam. Primjenom modela višestruke linearne regresije na transformiranim podacima dobiven je sljedeći model:

$$ln(y) = 5{,}331 - 0{,}01578x + 3{,}691I - 3{,}562N + 0{,}00003D$$

odnosno,

$$y = e^{5{,}331-0{,}01578x+3{,}691I-3{,}562N+0{,}00003D}$$

U Tablici 4.11 dani su rezultati višestruke linearne regresije između suficita gradova i gore navedenih nezavisnih varijabli. Prema koeficijentu uz transparentnost slijedi da povećanje transparentnosti dovodi do smanjenja suficita, suprotno istraživanjima navedenim u Poglavlju 1, no prema p-vrijednosti koeficijenta uz transparentnost slijedi da on nije statistički značajan. Koeficijenti uz kontrolne varijable također su neznačajni na svim razinama značajnosti.
Prilagođeni koeficijent determinacije R^2 = 0,003043, odnosno 0,3% prihoda po stanovniku gradova može se objasniti nezavisnim varijablama. F - statistika je jednaka 0,5963, a p - vrijednost modela 0,6664, stoga možemo zaključiti da model nije statistički značajan.

Koeficijent	**Procjena koeficijenta**	**Standardna devijacija**	**t - test statistika**	**p - vrijednost**
β_0	5,331	1,720	3,1	0,00272
β_1	-0,01578	0,1955	-0,081	0,93586
β_2	3,691	3,208	1,151	0,25348
β_3	-3,562	2,716	-1,312	0,19362
β_4	0,00003	0,00004	-0,670	0,50498

Tablica 4.11: Rezultati višestruke linearne regresije za treću hipotezu na gradovima

4.3.4 Testiranje Hipoteze 3 s kontrolnim varijablama na općinama

Kako bi se zadovoljile pretpostavke linearne regresije, proračunski suficit općina transformiran je koristeći logaritam. Primjenom modela višestruke linearne regresije na transformiranim podacima dobiven je sljedeći model:

$$ln(y) = 5{,}994 - 0{,}042x - 0{,}6680I + 1{,}473N + 0{,}00002D$$

odnosno,

$$y = e^{5{,}994-0{,}042x-0{,}6680I+1{,}473N+0{,}00002D}$$

U Tablici 4.12 dani su rezultati višestruke linearne regresije između suficita općina i gore navedenih nezavisnih varijabli. Prema koeficijentu uz transparentnost slijedi da povećanje transparentnosti dovodi do smanjenja suficita, suprotno istraživanjima navedenim u Poglavlju 1, no prema p-vrijednosti koeficijenta uz transparentnost slijedi da on nije statistički značajan. Koeficijenti uz kontrolne varijable također su neznačajni na svim razinama značajnosti.
Prilagođeni koeficijent determinacije $R^2 = 0{,}001089$, odnosno 0,1% suficita po stanovniku općina može se objasniti nezavisnim varijablama. Stoga možemo zaključiti da ne postoji veza između suficita po stanovniku općina i nezavisnih varijabli.

Koeficijent	**Procjena koeficijenta**	**Standardna devijacija**	**t - test statistika**	**p - vrijednost**
β_0	5,994	0,6747	8,883	<2e-16
β_1	-0,042	0,05861	-0,717	0,474
β_2	-0,6680	1,175	-0,569	0,570
β_3	1,473	1,007	1,463	0,144
β_4	0,00002	0,00002	0,929	0,354

Tablica 4.12: Rezultati višestruke linearne regresije za treću hipotezu na općinama

Treća hipoteza, koja tvrdi da povećanje proračunske transparentnosti doprinosi ostvarivanju proračunskog suficita, testirala se u ovom potpoglavlju. Kod gradova, rezultati jednostavne linearne regresije pokazali su da povećanje proračunske transparentnosti dovodi do povećanja suficita, ali da ta veza nije statistički značajna. Jednostavna linearna regresija kod općina i višestruka linearna regresija kod gradova i općina daje suprotni rezultat, odnosno da povećanje proračunske transparentnosti dovodi do smanjenja proračunskog suficita, ali opet, ta veza nije statistički značajna. Zaključak ovog potpoglavlja je da ne postoji veza između proračunske transparentnosti i prihoda po stanovniku hrvatskih općina i gradova.

4.4 Testiranje Hipoteze 4

Koristeći linearnu regresiju testirat će se četvrta hipoteza, odnosno doprinosi li veća transparentnost smanjenju proračunskog duga.

4.4.1 Testiranje Hipoteze 4 na gradovima

Kako bi se zadovoljile pretpostavke linearne regresije, proračunski dug kod gradova transformiran je koristeći treći korijen. Primjenom modela jednostavne linearne regresije na transformiranim podacima dobiven je sljedeći model:

$$\sqrt[3]{y} = 11{,}51 - 0{,}2501x$$

odnosno,

$$y = (11,51 - 0,2501x)^3$$

U tablici 4.13 dani su rezultati linearne regresije između proračunske transparentnosti i duga po stanovniku gradova. Prema koeficijentu uz transparentnost slijedi da povećanje transparentnosti dovodi do smanjenja duga po stanovniku gradova, suprotno istraživanjima navedenim u Poglavlju 1, no prema p-vrijednosti koeficijenta uz transparentnost slijedi da on nije statistički značajan.
Koeficijent determinacije R^2 = 0,002938, odnosno samo 0,3% javnog duga gradova može se objasniti s proračunskom transparentnosti. Stoga, možemo zaključiti da ne postoji veza između proračunske transparentnosti i duga po stanovniku gradova.

Koeficijent	**Procjena koeficijenta**	**Standardna devijacija**	**t - test statistika**	**p - vrijednost**
β_0	11,510	2,0648	5,574	1.91e-07
β_1	-0,2501	0,4475	-0,559	0,577

Tablica 4.13: Rezultati linearne regresije između proračunske transparentnosti i javnog duga po stanovniku gradova

4.4.2 Testiranje Hipoteze 4 na općinama

Kako bi se zadovoljile pretpostavke linearne regresije, proračunski dug kod općina transformiran je koristeći treći korijen. Primjenom modela jednostavne linearne regresije na transformiranim podacima dobiven je sljedeći model:

$$\sqrt[3]{y} = 9{,}60776 - 0{,}04405x$$

odnosno,

$$y = (9,60776 - 0,04405x)^3$$

U tablici 4.14 dani su rezultati linearne regresije između proračunske transparentnosti i duga po stanovniku općina. Prema koeficijentu uz transparentnost slijedi da povećanje transparentnosti dovodi do smanjenja javnog duga općina, suprotno istraživanjima navedenim u Poglavlju 1, no prema p-vrijednosti koeficijenta uz transparentnost slijedi da on nije statistički značajan.
Koeficijent determinacije R^2 = 0,00017, stoga, možemo zaključiti da ne postoji veza između proračunske transparentnosti i duga po stanovniku općina.

Koeficijent	Procjena koeficijenta	Standardna devijacija	t - test statistika	p - vrijednost
β_0	9,60776	0,89523	10,73	<2e-16
β_1	-0,04405	0,20965	-0,21	0,834

Tablica 4.14: Rezultati linearne regresije između proračunske transparentnosti i javnog duga po stanovniku općina

4.4.3 Testiranje Hipoteze 4 s kontrolnim varijablama na gradovima

Kako bi se zadovoljile pretpostavke linearne regresije, javni dug kod gradova transformiran je koristeći treći korijen. Primjenom modela višestruke linearne regresije na transformiranim podacima dobiven je sljedeći model:

$$\sqrt[3]{y} = 9{,}727 - 0{,}5926x + 7{,}160I - 0{,}8755N + 0{,}00003D$$

odnosno,

$$y = (9,727 - 0,5926x + 7,160I - 0,8755N + 0,00003D)^3$$

U Tablici 4.15 dani su rezultati višestruke linearne regresije između javnog duga gradova i gore navedenih nezavisnih varijabli. Prema koeficijentu uz transparentnost slijedi da povećanje transparentnosti dovodi do smanjenja javnog duga, kao i u istraživanjima navedenim u Poglavlju 1, no prema p-vrijednosti koeficijenta uz transparentnost slijedi da on nije statistički značajan. Koeficijenti uz kontrolne varijable također su neznačajni na svim razinama značajnosti.

Prilagođeni koeficijent determinacije R^2 = 0,0034, odnosno samo 0,3% javnog duga po stanovniku gradova može se objasniti nezavisnim varijablama. F - statistika je jednaka 1,091, a p - vrijednost modela 0,3562, stoga možemo zaključiti da model nije statistički značajan.

Koeficijent	Procjena koeficijenta	Standardna devijacija	t - test statistika	p - vrijednost
β_0	9,727	4,336	2,243	0,027
β_1	-0,5926	4,843	-1,223	0,224
β_2	7,160	7,828	0,915	0,363
β_3	-0,8755	7,947	-1,102	0,273
β_4	0,00003	0,0001	0,271	0,787

Tablica 4.15: Rezultati višestruke linearne regresije za četvrtu hipotezu na gradovima

4.4.4 Testiranje Hipoteze 4 s kontrolnim varijablama na općinama

Kako bi se zadovoljile pretpostavke linearne regresije, javni dug kod općina transformiran je koristeći treći korijen. Primjenom modela višestruke linearne regresije na transformiranim podacima dobiven je sljedeći model:

$$\sqrt[3]{y} = 9{,}552 - 0{,}1161x + 3{,}974I - 7{,}446N + 0{,}00002D$$

odnosno,

$$y = (9,552 - 0,1161x + 3,974I - 7,446N + 0,00002D)^3$$

U Tablici 4.16 dani su rezultati višestruke linearne regresije između javnog duga općina i gore navedenih nezavisnih varijabli. Prema koeficijentu uz transparentnost slijedi da povećanje transparentnosti dovodi do smanjenja javnog duga, kao i u istraživanjima navedenim u Poglavlju 1, no prema p-vrijednosti koeficijenta uz transparentnost slijedi da on nije statistički značajan. Koeficijent uz stopu nezaposlenosti statistički je značajan na razini značajnosti $\alpha = 5\%$, te njegovo povećavanje dovodi do smanjenja javnog duga. Prilagođeni koeficijent determinacije R^2 = 0,01818, odnosno 1,8% javnog duga po stanovniku općina može se objasniti nezavisnim varijablama. F - statistika je jednaka 2,204, a p -vrijednost modela 0,06902. Stoga možemo zaključiti model nije statistički značajan.

Koeficijent	**Procjena koeficijenta**	**Standardna devijacija**	**t - test statistika**	**p - vrijednost**
β_0	9,552	2,482	3,849	0,00015
β_1	-0,1161	2,091	-0,555	0,57926
β_2	3,974	4,278	0,929	0,35377
β_3	-7,446	3,740	-1,991	0,04757
β_4	0,00002	0,00006	-0,178	0,85888

Tablica 4.16: Rezultati višestruke linearne regresije za četvrtu hipotezu na općinama

Posljednja, četvrta hipoteza koja tvrdi da povećanje proračunske transparentnosti doprinosi smanjenju javnog duga testirala se u ovom potpoglavlju. Ona se testirala posebno po gradovima, posebno po općinama jednostavnom linearnom regresijom te uvođenjem kontrolnih varijabli višestrukom linearnom regresijom posebno po gradovima, posebno po općinama. Suprotno prethodnim radovima, jednostavnom linearnom regresijom dobiveno je da povećanje transparentnosti dovodi do povećanja proračunskog duga, ali da ta veza nije statistički značajna. Rezultati višestruke linearne regresije potvrdili su tvrdnju hipoteza, ali ni tu ta veza nije statistički značajna. Zaključak ovog potpoglavlja je da ne postoji veza između proračunske transparentnosti i javnog duga po stanovniku hrvatskih općina i gradova.

Poglavlje 5

Zaključak

Ovim radom pokušao se odrediti utjecaj proračunske transparentnosti na fiskalne rezultate hrvatskih općina i gradova. Cilj je bio utvrditi utječe li proračunska transparentnost na proračunski rashod po stanovniku, prihod po stanovniku, suficit po stanovniku i javni dug po stanovniku. Postavljene su hipoteze da transparentnost doprinosi povećanju proračunskog rashoda i prihoda, da doprinosi smanjenju duga te da doprinosi ostvarenju proračunskog suficita. Hipoteze su posebno testirane za gradove i posebno za općine.

Jednostavnom linearnom regresijom ustanovilo se da ne postoji statistički značajna veza između proračunske transparentnosti i proračunskog rashoda, prihoda, suficita i javnog duga po stanovniku. Uvođenje kontrolnih varijabli postotka korisnika interneta, stope nezaposlenosti i prosječnog dohotka u model višestruke regresije daje isti rezultat, odnosno transparentnost nema utjecaja na proračunski rashod, prihod, suficit i javni dug po stanovniku.

Razlog takvim rezultatima može se pronaći u samim podacima. Naime, čak 67% gradova ima transparentnost jednaku 5, dok njih 12% ima transparentnost manju ili jednaku 3. Dok frekvencija raste kako raste proračunska transparentnost gradova, takva ili suprotna veza nije postojala između proračunske transparentnosti i zavisnih varijabli gradova. Stoga je i bilo za očekivati da se neće moći pronaći veza između proračunske transparentnosti i fiskalnih rezultata hrvatskih općina i gradova. Slično, 47% općina ima najveću moguću transparentnost, dok 27% općina ima transparentnost manju ili jednaku 3. Također, kod deskriptivne statistike općina, najveće rashode po stanovniku, prihode po stanovniku i suficit po stanovniku imale su općine s proračunskom transparentnosti 0, dok je u hipotezama postavljeno da će se navedene zavisne varijable povećavati većom transparentnosti. Stoga se i na temelju tih podataka moglo očekivati da neće postojati veza između proračunske transparentnosti i fiskalnih rezultata hrvatskih općina i gradova.

Iako analiza u ovom radu zaključuje da proračunska transparentnost nema utjecaja na fiskalne rezultate hrvatskih općina i gradova, ona se itekako isplati. Njome se otvara prilika da se građani više uključe u donošenje odluka oko javnih financija. To bi tjeralo izvršitelje vlasti da odgovornije upravljaju javnim financijama te da ih koriste za javne interese.
U budućim istraživanjima mogla bi se provesti panel analiza na podacima dostupnim od 2015. da bi se utvrdilo kako promjena proračunske transparentnosti utječe na fiskalne rezultate kroz vrijeme.

Bibliografija

[1] A. Alesina, [et al.], *Budget institutions and fiscal performance in Latin America*, Journal of Development Economics dostupno na https://ssrn.com/abstract=1817191 (siječanj 2022)

[2] J. Alt i D. Lassen, *Fiscal transparency and fiscal policy outcomes in OECD Countries*, University of Copenhagen, dostupno na https://www.econstor.eu/bitstream/10419/82020/1/wp-03-02.pdf (siječanj 2022)

[3] J. Alt, D. Lassen i S. Rose, *The Causes of Fiscal Transparency: Evidence from the U.S. States*, International Monetary Fund, dostupno na https://core.ac.uk/download/pdf/7180013.pdf?repositoryId=153 (veljača 2022)

[4] J. Alt, D. Lassen i D. Skilling, *Fiscal Transparency, Gubernatorial Approval, and the Scale of Government: Evidence from the States*, dostupno na https://www.jstor.org/stable/40421462 (veljača 2022)

[5] F. Bastida i B. Benito, *Budget transparency, fiscal performance and political turonout: An international approach*, Public Administration Review (2009), 403-417

[6] M. Bronić, K. Ott, B. Stanić, F. Badovinac, *Ostvarenje proračuna općina, gradova i županija u 2020.*, dostupno na https://www.ijf.hr/upload/files/file/osvrti/123.pdf (srpanj 2021)

[7] M. Cimpoeru, V. Cimpoeru, *Budgetary Transparency – an Improving Factor for Corruption Control and Economic Performance*, dostupno na https://core.ac.uk/download/pdf/82777053.pdf (veljača 2022)

[8] Državni zavod za statistiku, *Popis stanovništva 2011. godine*, dostupno na https://www.dzs.hr/hrv/censuses/census2011/results/censustabsxls.htm (kolovoz 2021)

[9] T. Gérard, N. Ngangué, *Does Fiscal Illusion Impact Budget Policy? A Panel Data Analysis*, International Journal of Economics and Financial Issues (2014), 240–248

[10] F. Hameed, *Fiscal transparency and economic outcomes*, dostupno na https://ssrn.com/abstract=888094 (veljača 2022)

[11] M. Huzak, *Statistički praktikum – linearna regresija*, dostupno na https://web.math.pmf.unizg.hr/nastava/statpr/files/linearna.pdf (veljača 2022)

[12] G. Kopits i J. Craig, *Transparency in government operations*, Occasional Paper 158

[13] Ministarstvo financija Republike Hrvatske, *Financijski izvještaji JLP(R)S*, dostupno na https://mfin.gov.hr/istaknute-teme/lokalna-samouprava/financijski-izvjestaji-jlp-r-s/203 (srpanj 2021)

[14] C.Perez R.Bolivar, L. Hernandez, *e-Government process and incentives for online public financial information*, Emeral Insight, 2008.

[15] S. J. Piotrowski i G. Van Ryzin, *Citizen Attitudes Toward Transparency in Local Government*, The American Review of Public Administration, 2007. dostupno na https://journals.sagepub.com/doi/10.1177/0275074006296777 (srpanj 2020)

[16] Regionalni razvoj, *Indeks razvijenosti*, dostupno na http://regionalni.weebly.com/indeksrazvijenosti.html (kolovoz 2021)

[17] A. C. Rencher i G. B. Schaalje, *Linear models in statistics*, Department of Statistics, Brigham Young University, Provo, Utah, 2007.

[18] P. de Renzio, J. Wehner, *The Impacts of Fiscal Openness: A Review of the Evidence*, dostupno na https://personal.lse.ac.uk/wehner/openness.pdf (veljača 2022)

[19] E. Stein, E. Talvi i A. Grisanti, *Institutional Arrangements and Fiscal Performance: The Latin American Experience*, dostupno na https://www.nber.org/system/files/chapters/c8025/c8025.pdf (veljača 2022)

[20] Proračunska transparentnost županija, gradova i općina: studeni 2020. – travanj 2021., dostupno na https://www.ijf.hr/upload/files/file/osvrti/119.pdf (srpanj 2022)

[21] A. F. Tavares, N. F. da Cruz, *Explaining the transparency of local government websites through a political market framework*, dostupno na https://www.sciencedirect.com/science/article/abs/pii/S0740624X17302617 (svibanj 2022)

[22] Uredba o indeksu razvijenosti, dostupno na https://narodne-novine.nn.hr/clanci/sluzbeni/2017121313014.html (siječanj 2022)

[23] C. Vicente, B. Benito i F. Bastida, Transparency and Political Cycles at municipal level, Univesity of Murcia, dostupno na https://onlinelibrary.wiley.com/toc/16626370/2013/19/2 (siječanj 2022)

Sažetak

U ovom radu cilj je bio utvrditi utjecaj proračunske transparentnosti na fiskalne rezultate hrvatskih općina i gradova. Nezavisna varijabla bila je transparentnost proračuna, a zavisne varijable bile su proračunski rashod po stanovniku, proračunski prihod po stanovniku, proračunski suficit po stanovniku te javni dug po stanovniku. Postavljene su četiri hipoteze te su prikupljeni podaci o 428 općina i 128 gradova u Republici Hrvatskoj za 2020. godinu. Modelom jednostavne linearne regresije utvrđeno je da ne postoji statistički značajna veza između proračunske transparentnosti i zavisnih varijabli ni za općine ni za gradove. Korištenjem višestruke linearne regresije uz kontrolne varijable postotak korisnika interneta, stope nezaposlenosti i prosječnog dohotka dobiveni su isti rezultati, odnosno proračunska transparentnost nema statistički značajan utjecaj na zavisne varijable, ni kod općina, ni kod gradova.

Summary

In this paper, the goal was to determine the impact of budget transparency on the fiscal results of Croatian municipalities and cities. The independent variable was budget transparency, and the dependent variables were budget expenditure per capita, budget revenue per capita, budget surplus per capita and public debt per capita. Four hypotheses were set and data on 428 municipalities and 128 cities in the Republic of Croatia for the year 2020 was collected. Using a simple linear regression model, it could be concluded that there is no statistically significant relationship between budget transparency and dependent variables for either municipalities or cities. By using multiple linear regression with the control variables percentage of internet users, unemployment rate and average income, the same results were obtained, i.e. budget transparency has no statistically significant impact on the dependent variables neither for municipalities nor for cities.

Životopis

Lucija Marijan rođena je 10. ožujka 1997. u Zagrebu. Nakon završene osnovne škole Eugena Kumičića u Velikoj Gorici pohađa XV. gimnaziju u Zagrebu. Završetkom srednjoškolskog obrazovanja upisuje preddiplomski studij Matematike na Prirodoslovno-matematičkom fakultetu Sveučilišta u Zagrebu. Nakon završetka preddiplomskog studija, na istom fakultetu upisuje diplomski studij, smjer Financijska i poslovna matematika