

Korekcija anketnih podataka s pomoću poreznih podataka: metoda i primjena

Anketni podaci

- Anketni podaci su još od 1970-ih neizostavan izvor podataka za istraživanja u ekonomiji, sociologiji, politologiji...
- Na nacionalno reprezentativnom uzorku kućanstava prikupljaju se podaci o izvorima dohotka, socio-ekonomskim i demografskim obilježjima kućanstava i njihovih članova, izdacima za potrošnju...
- Primjeri kod nas: Anketa o potrošnji kućanstava (APK), Anketa o dohotku stanovništva (ADS), Anketa o radnoj snazi (ARS), European Union Statistics on Income and Living Conditions (EU-SILC)...
- Anketni podaci korišteni u ovom projektu: APK i EU-SILC
- Za mjerenje i dekompozicije nejednakosti, mikrosimulacije poreza i socijalnih naknada, mjerenje progresivnosti poreza...

Anketni podaci često ne odražavaju dobro stvarnu distribuciju dohotka

- Brojna istraživanja pokazuju da **dohoci na vrhu i na dnu distribucije** ne reprezentiraju dobro stvarno ili “stvarno” stanje (utoliko ukoliko ga je moguće utvrditi iz drugih izvora)
- Problem može nastati **već kod uzorkovanja**: manja vjerojatnost odabira u uzorak (namjera anketiranja) društvenih skupina s obilježjima koja koreliraju s dohotkom
- **Čak i ako je vjerojatnost da se nađu u uzorku jednaka za sve**, problemi mogu nastati jer
 - **Relativno mali uzorci** ne mogu zahvatiti neke male skupine, npr. iznimno bogate ili siromašne
 - **Spremnost na sudjelovanje u anketi** može varirati s dohotkom
 - **Spremnost da se odgovori na određena pitanja** (npr., izvori dohotka) može varirati s dohotkom
 - **Veličina podcjenjivanja dohotka** može varirati s dohotkom
 - Na vrhu distribucije može biti primijenjen tzv. **top-coding**

Anketni podaci o dohotku se mogu korigirati kako bi odgovarali stvarnima

- Anketni podaci o dohotku se mogu korigirati **ako imamo neki izvor “istine”** izvan same ankete, odnosno **“dovoljno vjerodostojne” informacije** o stvarnim dohocima
- **Administrativni podaci** o dohocima koje prikuplja država za potrebe oporezivanja – **porezni podaci** – smatraju se relativno vjerodostojnim, u usporedbi s anketnim podacima
- Korištenje poreznih podataka za analize nejednakosti doživjelo procvat u zadnjih 25 godina (Tony Atkinson, Thomas Piketty; *World Inequality Database*)
- Postoje **metode korekcije dohodaka u anketama s pomoću poreznih podataka**
- **Zašto korigirati anketne podatke ako imamo porezne?** Porezni podaci obično sadrže vrlo malo ne-dohodovnih varijabli

MARKO LEDIĆ, IVICA RUBIL, IVICA URBAN

**Using tax records to correct for under-representation of top
income sources in surveys**

Fiscal Studies

2024

Motivacija

- Dosadašnje korekcije anketnih podataka na vrhu distribucije – radi nepristranog **mjerenja nejednakosti**
- Postojeće metode korekcije fokusirane na korekciju **ukupnog** dohotka (npr., ukupni bruto dohodak)
- Međutim, ispravni dohoci potrebni su i u **drugim primjenama**:
 - **Mikrosimulacija** poreza i socijalnih naknada
 - **Dekompozicija** nejednakosti ukupnoga dohotka na doprinose pojedinih izvora dohotka
- Za nepristrane rezultate u tim primjenama, potrebna korekcija ne samo ukupnog dohotku, nego i **dohodaka iz pojedinih izvora**, dakle **strukture** ukupnog dohotka
 - Važno u **mikrosimulacijama** ako su dohoci iz različitih izvora oporezovani različitim stopama (u većini slučajeva)
 - Važno u **dekompozicijama nejednakosti** jer doprinosi pojedinih izvora ovise o njihovoj distribuciji i udjelima u ukupnom dohotku

Što radimo; doprinos

- **Korekcija dohodaka na vrhu distribucije u anketi EU-SILC** (verzija za EUROMOD)
- **Korekcija s pomoću podataka Porezne uprave:**
 - podaci prikupljeni na temelju obrasca JOPPD i drugih obrazaca za prijavu dohodaka
- **Unapređenja recentne metode korekcije** – Blanchet, Flores i Morgan, 2022., *Journal of Economic Inequality* – “**BFM metoda**”
- **Zašto unaprjeđenja:** BFM metoda je fokusirana na korekciju ukupnog dohotka; korekcija pojedinih izvora dohotka načelno moguća, ali ne funkcionira ako struktura ukupnog dohotka u anketi previše odstupa od one u poreznim podacima
- Unaprjeđena BFM metoda **uspijeva istodobno korigirati ukupni dohodak i približiti njegovu strukturu onoj u poreznim podacima**

BFM metoda i naša unaprjeđenja

BFM metoda

- Pretpostavlja da u anketi **fali dohotka na vrhu** distribucije
- **Reponderiranje (reweighting)** – povećanje (smanjenje) anketnih pondera onih s višim (nižim) dohocima
- **Zamjena (replacement)** visokih anketnih dohodaka još višim dohocima iz poreznih podataka – spajanje anketne i porezne distribucije iznad određene razine dohotka (**merging point**)
- Ako dohotka iz nekih izvora previše fali, struktura po izvorima u anketnim podacima se ne može popraviti

Naša dva unaprjeđenja / nadopune

- **Pred-korekcija + BFM metoda**: “presađivanje” izvora dohotka kojih previše fali iz poreznih podataka u anketne, **prije** BFM metode
- **BFM metoda + post-korekcija**: “nametanje” strukture dohotka *iznad merging pointa* iz poreznih u anketne podatke, **nakon** BFM metode

U anketi fali 11.3% ukupnog dohotka iz poreznih podataka; najviše zbog manjka dohotka od zaposlenosti i dohotka od kapitala, i to na vrhu distribucije

Ukupni dohodak = “fiscal income” = ukupni bruto dohodak

$$\text{All} = \text{p0-p40} + \text{p40-p60} + \text{p60-p80} + \text{p80-p90} + \text{p90-p95} + \text{p95-p99} + \text{p99-p100}$$

A. Per cent of aggregate fiscal income in TD

Fiscal income	-11.3	-1.5	-0.8	1.4	1.3	0.1	-3.8	-8.1
=								
Employment income	-4.8	-0.3	-0.8	2.1	0.7	-0.2	-2.3	-4.0
+								
Self-employment income	-0.4	0.0	0.1	0.6	0.2	-0.3	-0.7	-0.3
+								
Pension income	0.9	-0.8	0.3	-0.7	0.9	1.0	0.3	-0.1
+								
Contractual work income	-1.3	0.0	0.0	-0.2	-0.1	-0.1	-0.3	-0.4
+								
Property income	-1.4	-0.1	-0.2	-0.2	-0.2	-0.1	-0.4	-0.2
+								
Capital income	-4.3	-0.2	-0.1	-0.2	-0.2	-0.1	-0.4	-3.1

Od originalnih do korigiranih anketnih podataka

U originalnoj anketi
fali 11.3% ukupnog
bruto dohotka iz
poreznih podataka...

... najviše zbog toga što fali
7.1% dohoka od zaposlenosti i
skoro sav dohodak od kapitala

Agregatni iznos dohotka u anketi kao % iznosa u poreznim podacima

	OR	BFM	PR	PR-BFM
Fiscal income	-11.3	-0.6	-7.0	1.6
Employment income	-7.1	7.2	-7.1	0.7
Self-employment income	-10.6	-0.1	-10.6	-1.4
Pension income	4.5	4.7	4.5	5.5
Contractual work income	-63.1	-43.1	-0.2	28.6
Property income	-56.2	-48.5	-3.6	7.1
Capital income	-97.5	-97.0	-57.7	-15.6

Od originalnih do korigiranih anketnih podataka

BFM metoda
uspjeva korigirati
ukupni dohodak...

... ali sada dohotka od rada ima
previše, dok je dohodak od kapitala
ostao gotovo isti kao prije korekcije

Agregatni iznos dohotka u anketi kao % iznosa u poreznim podacima

	OR	BFM	PR	PR-BFM
Fiscal income	-11.3	-0.6	-7.0	1.6
Employment income	-7.1	7.2	-7.1	0.7
Self-employment income	-10.6	-0.1	-10.6	-1.4
Pension income	4.5	4.7	4.5	5.5
Contractual work income	-63.1	-43.1	-0.2	28.6
Property income	-56.2	-48.5	-3.6	7.1
Capital income	-97.5	-97.0	-57.7	-15.6

Od originalnih do korigiranih anketnih podataka

Pred-korekcijom djelomično korigiramo dohotke od kapitala i imovine i drugi dohodak...

... ali sama pred-korekcija može samo djelomično korigirati ukupni dohodak; potrebno dodati BFM metodu

Agregatni iznos dohotka u anketi kao % iznosa u poreznim podacima

	OR	BFM	PR	PR-BFM
Fiscal income	-11.3	-0.6	-7.0	1.6
Employment income	-7.1	7.2	-7.1	0.7
Self-employment income	-10.6	-0.1	-10.6	-1.4
Pension income	4.5	4.7	4.5	5.5
Contractual work income	-63.1	-43.1	-0.2	28.6
Property income	-56.2	-48.5	-3.6	7.1
Capital income	-97.5	-97.0	-57.7	-15.6

Od originalnih do korigiranih anketnih podataka

Kombinacija pred-korekcije i BFM metode uspijeva korigirati ukupni dohodak...

... i istodobno značajno (ali ne savršeno) korigira izvore dohotka u odnosu na originalnu anketu

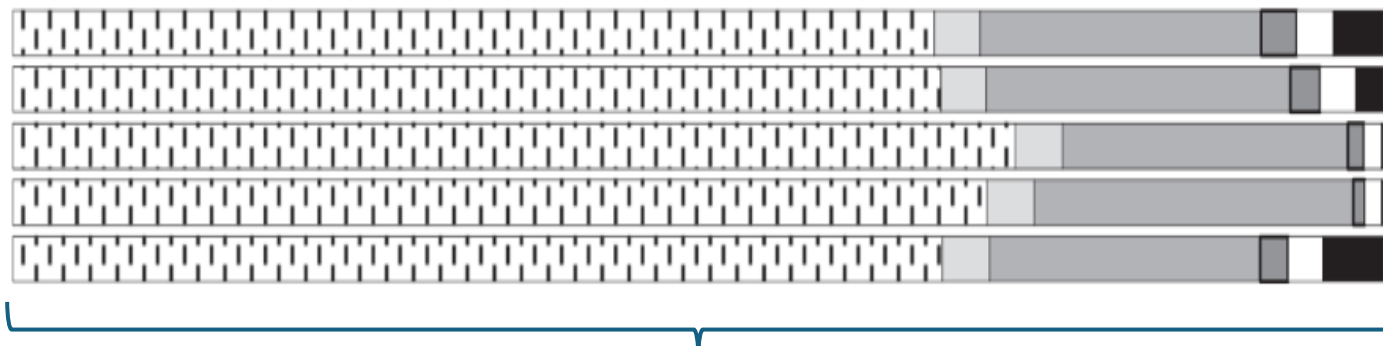
Agregatni iznos dohotka u anketi kao % iznosa u poreznim podacima

	OR	BFM	PR	PR-BFM
Fiscal income	-11.3	-0.6	-7.0	1.6
Employment income	-7.1	7.2	-7.1	0.7
Self-employment income	-10.6	-0.1	-10.6	-1.4
Pension income	4.5	4.7	4.5	5.5
Contractual work income	-63.1	-43.1	-0.2	28.6
Property income	-56.2	-48.5	-3.6	7.1
Capital income	-97.5	-97.0	-57.7	-15.6

Značajno poboljšanje strukture ukupnog dohotka po izvorima

□ Employment □ Self-employment □ Pension □ Contractual work □ Property ■ Capital

Pred-kor. + BFM met.
Samo pred-korekcija
Samo BFM metoda
Originalni anketni
Porezni podaci



Ukupni dohodak = 100%

SLAVKO BEZEREDI, NORA MUSTAĆ, IVICA URBAN

**Dekompozicije nejednakosti prema izvorima dohotka:
što nam govore anketni podaci korigirani uz pomoć poreznih
podataka**

Ekonomska misao i praksa

2023.

Sažetak

- Dekompozicija nejednakosti ukupnoga raspoloživog dohotka na sastavnice koje pokazuju doprinose pojedinih izvora dohotka
- Anketni podaci iz EU-SILC-a (EUROMOD verzija)
- Od bruto dohotka do raspoloživog dohotka s pomoću mikrosimulacijskog modela EUROMOD
- **Novina:** anketni podaci korigirani (pretkorekcija + BFM) s pomoću podataka Porezne uprave kako bi gornji dio distribucije bolje odražavao stvarnu distribuciju
- **Rezultati:** Korištenje originalnih, nekorigiranih podataka uzrokuje pristranost u procjenama razine ukupne nejednakosti i razmjernim doprinosima pojedinih izvora dohotka

Ginijev koef. prema **korigiranim** podacima
(0,332) znatno je veći nego prema **originalnim**
 podacima **(0,290)**...

	Prirodna dekompozicija – nominalni doprinosi		Prirodna dekompozicija – razmjerni doprinosi	
	$C_k^R(G)$		$100 \cdot c_k^R(G)\%$	
	HR	HR*	HR	HR*
Ukupni dohodak	0,290	0,332	100,0	100,0
Dohodak od plaće	0,346	0,349	119,5	105,1
Dohodak od samostalne djelatnosti	0,036	0,037	12,4	11,1
Dohodak od kapitala i imovine	0,008	0,056	2,8	16,9
Ostali privatni dohodak	0,000	0,000	0,0	0,0
Javne mirovine	0,029	0,020	9,9	6,0
Socijalne naknade s cenzusom	-0,010	-0,008	-3,3	-2,4
Socijalne naknade bez cenzusa	0,002	0,001	0,8	0,3
Doprinosi za obvezna osiguranja	-0,076	-0,070	-26,3	-21,2
Porezi na dohodak i imovinu	-0,046	-0,053	-15,9	-15,9

Ginijev koef. prema **korigiranim** podacima
(0,332) znatno je veći nego prema **originalnim**
 podacima **(0,290)**...

	Prirodna dekompozicija – nominalni doprinosi		Prirodna dekompozicija – razmjerni doprinosi	
	$C_k^R(G)$		$100 \cdot c_k^R(G)\%$	
	HR	HR*	HR	HR*
Ukupni dohodak	0,290	0,332	100,0	100,0
Dohodak od plaće	0,346	0,349	119,5	105,1
Dohodak od samostalne djelatnosti	0,036	0,037	12,4	11,1
Dohodak od kapitala i imovine	0,008	0,056	2,8	16,9
Ostali privatni dohodak	0,000	0,000	0,0	0,0
Javne mirovine	0,029	0,020	9,9	6,0
Socijalne naknade s cenzusom	-0,010	-0,008	-3,3	-2,4
Socijalne naknade bez cenzusa	0,002	0,001	0,8	0,3
Doprinosi za obvezna osiguranja	-0,076	-0,070	-26,3	-21,2
Porezi na dohodak i imovinu	-0,046	-0,053	-15,9	-15,9

... uglavnom zbog
povećanja doprinosa
dohotka od kapitala i
imovine

Razmjerni (%-tni) doprinosi dohotka od kapitala i imovine veći je prema **korigiranim** podacima nego prema **originalnim** podacima...

	Prirodna dekompozicija – nominalni doprinosi		Prirodna dekompozicija – razmjerni doprinosi	
	$C_k^R(G)$		$100 \cdot c_k^R(G)\%$	
	HR	HR*	HR	HR*
Ukupni dohodak	0,290	0,332	100,0	100,0
Dohodak od plaće	0,346	0,349	119,5	105,1
Dohodak od samostalne djelatnosti	0,036	0,037	12,4	11,1
Dohodak od kapitala i imovine	0,008	0,056	2,8	16,9
Ostali privatni dohodak	0,000	0,000	0,0	0,0
Javne mirovine	0,029	0,020	9,9	6,0
Socijalne naknade s cenzusom	-0,010	-0,008	-3,3	-2,4
Socijalne naknade bez cenzusa	0,002	0,001	0,8	0,3
Doprinosi za obvezna osiguranja	-0,076	-0,070	-26,3	-21,2
Porezi na dohodak i imovinu	-0,046	-0,053	-15,9	-15,9

Razmjerni (%-tni) doprinosi dohotka od kapitala i imovine veći je prema **korigiranim** podacima nego prema **originalnim** podacima...

	Prirodna dekompozicija – nominalni doprinosi		Prirodna dekompozicija – razmjerni doprinosi	
	$C_k^R(G)$		$100 \cdot c_k^R(G)\%$	
	HR	HR*	HR	HR*
Ukupni dohodak	0,290	0,332	100,0	100,0
Dohodak od plaće	0,346	0,349	119,5	105,1
Dohodak od samostalne djelatnosti	0,036	0,037	12,4	11,1
Dohodak od kapitala i imovine	0,008	0,056	2,8	16,9
Ostali privatni dohodak	0,000	0,000	0,0	0,0
Javne mirovine	0,029	0,020	9,9	6,0
Socijalne naknade s cenzusom	-0,010	-0,008	-3,3	-2,4
Socijalne naknade bez cenzusa	0,002	0,001	0,8	0,3
Doprinosi za obvezna osiguranja	-0,076	-0,070	-26,3	-21,2
Porezi na dohodak i imovinu	-0,046	-0,053	-15,9	-15,9

... uglavnom nauštrb doprinosa dohotka od plaće

Zašto?

- S **korekcijom** se **povećava udio dohotka od kapitala i imovine u ukupnom dohotku**, dok se **udio ostalih izvora dohotku smanjuje**
- Dohodak od kapitala i imovine je koncentriraniji na vrhu nego dohodak iz ostalih izvora (osim poreza na dohodak i imovinu)
- K tome, s korekcijom koncentracija dohotka od kapitala i imovine raste (znatno više nego koncentracija ostalih izvora)

	Prirodna dekompozicija – nominalni doprinosi		Prirodna dekompozicija – razmjerni doprinosi		Udio u ukupnom dohotku		Ginijev indeks koncentracije	
	$C_k^R(G)$		$100 \cdot c_k^R(G)\%$		φ_k		$D^y(x_k)$	
	HR	HR*	HR	HR*	HR	HR*	HR	HR*
Ukupni dohodak	0,290	0,332	100,0	100,0	1,000	1,000	0,290	0,332
Dohodak od plaće	0,346	0,349	119,5	105,1	0,844	0,800	0,410	0,437
Dohodak od samostalne djelatnosti	0,036	0,037	12,4	11,1	0,100	0,100	0,361	0,370
Dohodak od kapitala i imovine	0,008	0,056	2,8	16,9	0,015	0,080	0,539	0,703
Ostali privatni dohodak	0,000	0,000	0,0	0,0	0,011	0,010	-0,007	-0,009
Javne mirovine	0,029	0,020	9,9	6,0	0,237	0,207	0,122	0,097
Socijalne naknade s cenzusom	-0,010	-0,008	-3,3	-2,4	0,015	0,013	-0,618	-0,620
Socijalne naknade bez cenzusa	0,002	0,001	0,8	0,3	0,026	0,021	0,096	0,055
Doprinosi za obvezna osiguranja	-0,076	-0,070	-26,3	-21,2	-0,187	-0,163	0,408	0,431
Porezi na dohodak i imovinu	-0,046	-0,053	-15,9	-15,9	-0,061	-0,068	0,754	0,775

Zašto?

- S korekcijom se **povećava udio dohotka od kapitala i imovine u ukupnom dohotku**, dok se **udio ostalih izvora dohotku smanjuje**
- Dohodak od kapitala i imovine je **koncentriraniji** na vrhu nego dohodak iz ostalih izvora (osim poreza na dohodak i imovinu)
- K tome, s korekcijom koncentracija dohotka od kapitala i imovine raste (znatno više nego koncentracija ostalih izvora)

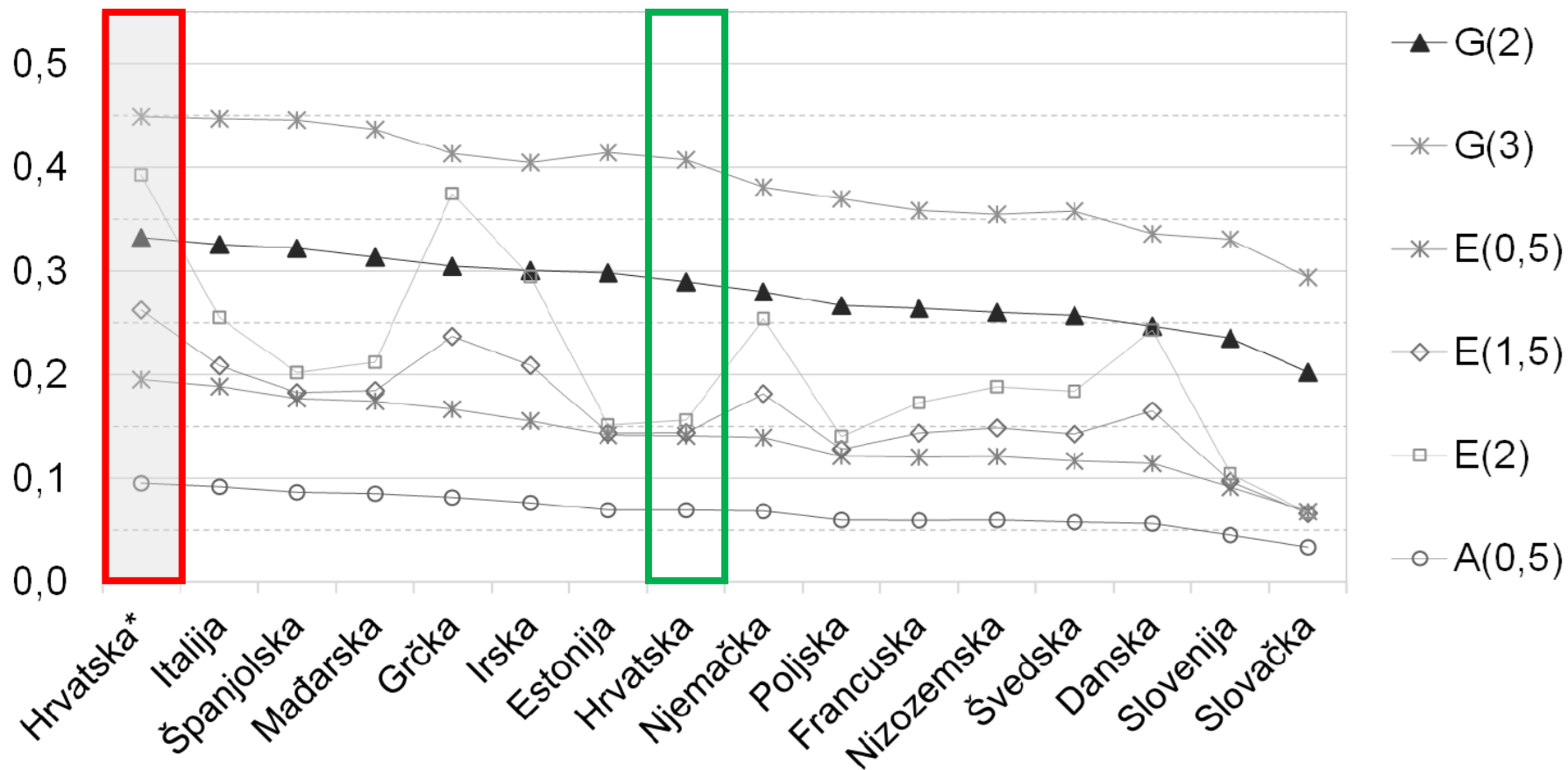
	Prirodna dekompozicija – nominalni doprinosi		Prirodna dekompozicija – razmjerni doprinosi		Udio u ukupnom dohotku		Ginijev indeks koncentracije	
	$C_k^R(G)$		$100 \cdot c_k^R(G)\%$		φ_k		$D^y(x_k)$	
	HR	HR*	HR	HR*	HR	HR*	HR	HR*
Ukupni dohodak	0,290	0,332	100,0	100,0	1,000	1,000	0,290	0,332
Dohodak od plaće	0,346	0,349	119,5	105,1	0,844	0,800	0,410	0,437
Dohodak od samostalne djelatnosti	0,036	0,037	12,4	11,1	0,100	0,100	0,361	0,370
Dohodak od kapitala i imovine	0,008	0,056	2,8	16,9	0,015	0,080	0,539	0,703
Ostali privatni dohodak	0,000	0,000	0,0	0,0	0,011	0,010	-0,007	-0,009
Javne mirovine	0,029	0,020	9,9	6,0	0,237	0,207	0,122	0,097
Socijalne naknade s cenzusom	-0,010	-0,008	-3,3	-2,4	0,015	0,013	-0,618	-0,620
Socijalne naknade bez cenzusa	0,002	0,001	0,8	0,3	0,026	0,021	0,096	0,055
Doprinosi za obvezna osiguranja	-0,076	-0,070	-26,3	-21,2	-0,187	-0,163	0,408	0,431
Porezi na dohodak i imovinu	-0,046	-0,053	-15,9	-15,9	-0,061	-0,068	0,754	0,775

Zašto?

- S korekcijom se **povećava udio dohotka od kapitala i imovine u ukupnom dohotku**, dok se **udio ostalih izvora dohotku smanjuje**
- Dohodak od kapitala i imovine je **koncentriraniji na vrhu** nego dohodak iz ostalih izvora (osim poreza na dohodak i imovinu)
- K tome, s **korekcijom koncentracija dohotka od kapitala i imovine raste** (znatno više nego koncentracija ostalih izvora)

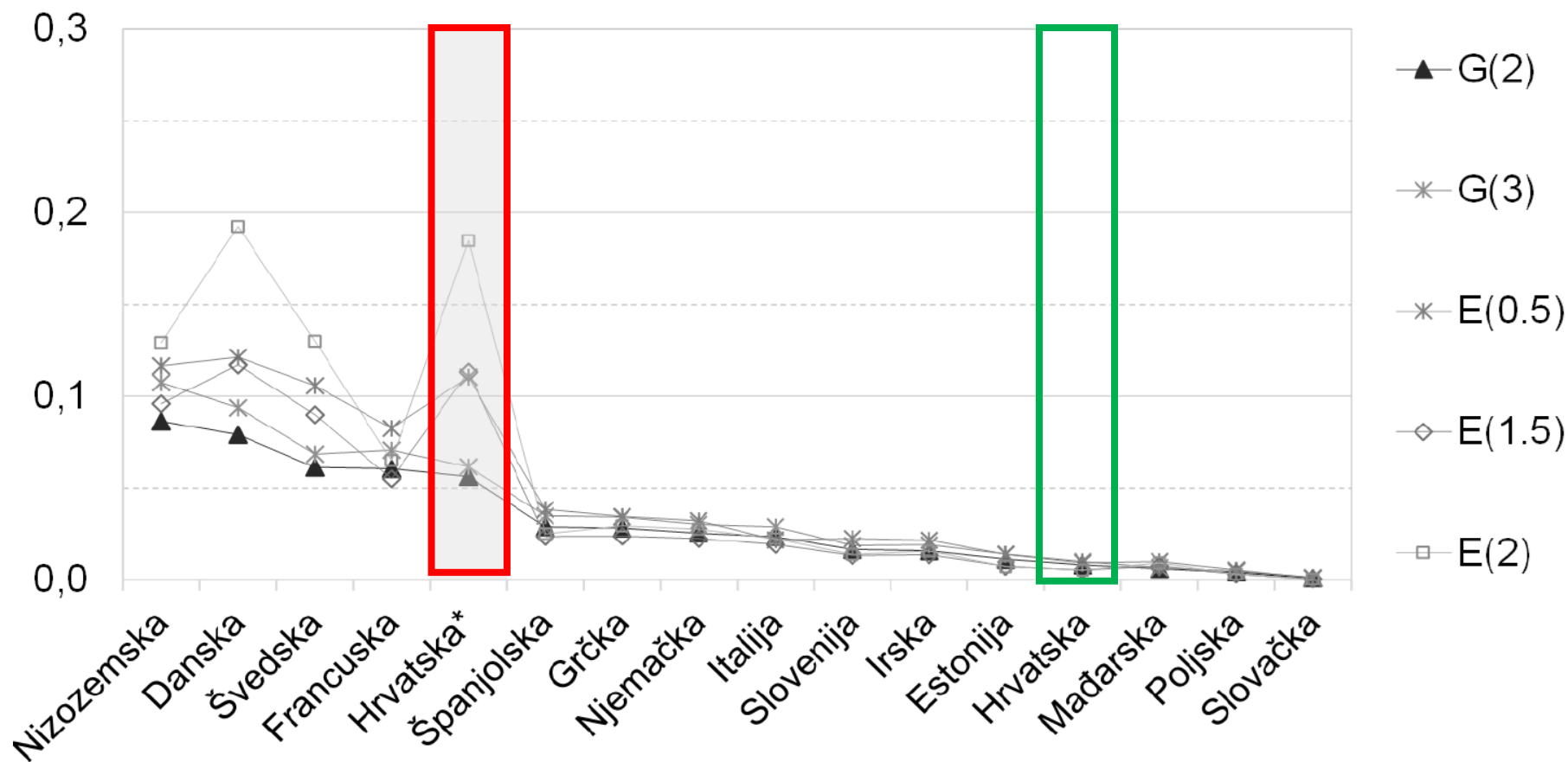
	Prirodna dekompozicija – nominalni doprinosi		Prirodna dekompozicija – razmjerni doprinosi		Udio u ukupnom dohotku		Ginijev indeks koncentracije	
	$C_k^R(G)$		$100 \cdot c_k^R(G)\%$		φ_k		$D^y(x_k)$	
	HR	HR*	HR	HR*	HR	HR*	HR	HR*
Ukupni dohodak	0,290	0,332	100,0	100,0	1,000	1,000	0,290	0,332
Dohodak od plaće	0,346	0,349	119,5	105,1	0,844	0,800	0,410	0,437
Dohodak od samostalne djelatnosti	0,036	0,037	12,4	11,1	0,100	0,100	0,361	0,370
Dohodak od kapitala i imovine	0,008	0,056	2,8	16,9	0,015	0,080	0,539	0,703
Ostali privatni dohodak	0,000	0,000	0,0	0,0	0,011	0,010	-0,007	-0,009
Javne mirovine	0,029	0,020	9,9	6,0	0,237	0,207	0,122	0,097
Socijalne naknade s cenzusom	-0,010	-0,008	-3,3	-2,4	0,015	0,013	-0,618	-0,620
Socijalne naknade bez cenzusa	0,002	0,001	0,8	0,3	0,026	0,021	0,096	0,055
Doprinosi za obvezna osiguranja	-0,076	-0,070	-26,3	-21,2	-0,187	-0,163	0,408	0,431
Porezi na dohodak i imovinu	-0,046	-0,053	-15,9	-15,9	-0,061	-0,068	0,754	0,775

Prema originalnim (korigiranim) podacima, Hrvatska je na sredini (vrhu) ljestvice po nejednakosti u usporedbi s odabranim zemljama



Podaci za druge zemlje nisu korigirani; neke zemlje EU-SILC popunjavaju arministrativnim podacima.

Prema **originalnim** (**korigiranim**) podacima, u Hrvatskoj je proporcionalni doprinos dohotka od kapitala ukupnoj nejedakosti među **nižima** (**višima**) u usporedbi s odabranim zemljama



Podaci za druge zemlje nisu korigirani; neke zemlje EU-SILC popunjavaju arministrativnim podacima.