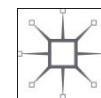


PRIMIJENJENA EKONOMETRIJA

MULTIKOLINEARNOST



MULTIKOLINEARNOST

1. Savršena multikolinearnost
2. Posledice postojanja savršene multikolinearnosti
3. Približna multikolinearnost
4. Posledice postojanja približne multikolinearnosti
5. Načini za otkrivanje postojanja multikolinearnosti
6. Rešavanje problema postojanja multikolinearnosti



PRIMIJENJENA EKONOMETRIJA

Multikolinearnost

Jedna od pretpostavki KLRM navodi da između objašnjavajućih promjenljivih ne postoji tačna linearna zavisnost.

Dakle, kada su objašnjavajuće promjenljive visoko korelisane (koeficijent korelacije uzima vrednost ili približno jednaku -1 ili 1) tada nastaje problem štetne multikolinearnosti.



PRIMIJENJENA EKONOMETRIJA

Savršena multikolinearnost

- Kada među objašnjavajućim promenljivim postoji perfektna linearna veza
- Pretpostavimo da imamo sledeći model:

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + e$$

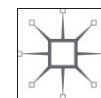
gde su uzoračke vrednosti za X_2 i X_3 :

X_2	1	2	3	4	5	6
X_3	2	4	6	8	10	12



Savršena multikolinearnost

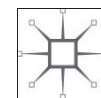
- Zaključujemo da je $X_3 = 2X_2$
- Dakle, iako se činilo da imamo dve promenljive, zapravo postoji samo jedna, jer je promenljiva X_2 tačna linearna funkcija promenljive X_3 . Prema tome, varijable X_2 i X_3 su perfektno korelisane.



PRIMIJENJENA EKONOMETRIJA

Posledice postojanja savršene multikolinearnosti

- U situaciji kada postoji savršena multikolinearnost model ne može da se oceni.
- Ne može da se razdvoji pojedinačni uticaj objašnjavajućih promenljivih



Približna multikolinearnost

- Približna multikolinearnost nastaje kada među objašnjavajućim promenljivim postoji korelacija, ali ona nije savršena.
- Odnos između dve objašnjavajuće promenljive u tom slučaju može da se prikaže sledećom relacijom:

$$X_3 = X_2 + v$$

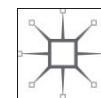
gde je v slučajna greška tačne linearne korelacije.



PRIMIJENJENA EKONOMETRIJA

Posledice postojanja približne multikolinearnosti

- Ocene parametara ostaju nepristrasne
- Ocene regresionih parametara su neprecizne u smislu visokih standardnih grešaka ocena.
- t-odnosi su niski i mogu dovesti do pogrešnog statističkog zaključka
- Visoka vrednost koeficijenta determinacije je praćena niskim t-odnosima



PRIMIJENJENA EKONOMETRIJA

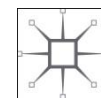
Posledice postojanja približne multikolinearnosti

Varijanse ocena parametara uz objašnjavaju'e
promenljive X_2 i X_3 su:

$$\text{var}(\hat{\beta}_2) = \frac{\sigma^2}{\sum (X_2 - \bar{X}_2)^2 (1 - r^2)}$$

$$\text{var}(\hat{\beta}_3) = \frac{\sigma^2}{\sum (X_3 - \bar{X}_3)^2 (1 - r^2)}$$

gde je r^2 kvadrat koeficijenta korelacije varijabli X_2 i X_3 .



PRIMIJENJENA EKONOMETRIJA

Posledice postojanja približne multikolinearnosti

Ako imamo više od dve objašnjavajuće promenljive, tada je varijansa ocean parametara uz ove promenljive:

$$\text{var}(\hat{\beta}_j) = \frac{\sigma^2}{\sum (X_2 - \bar{X}_2)^2 (1 - R_j^2)}$$

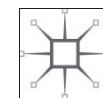
$$\text{var}(\hat{\beta}_3) = \frac{\sigma^2}{\sum (X_3 - \bar{X}_3)^2} \frac{1}{(1 - R_j^2)}$$



PRIMIJENJENA EKONOMETRIJA

Inflatorni faktor varijanse

R^2_j	VIF _j
0	1
0.5	2
0.8	5
0.9	10
0.95	20
0.975	40
0.99	100
0.995	200
0.999	1000



Inflatorni faktor varijanse

- Kada je VIF vrednost veća od 10, tada postoji problem štetne multikolinearnosti.
- To je situacija u kojoj je $R^2_j > 0.9$



PRIMIJENJENA EKONOMETRIJA

Ispitivanje postojanja multikolinearnosti

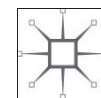
- Najjedostavniji način jeste analiza matrice koeficijenata korelacije objašnjavajućih promenljivih.
- U slučaju postojanja samo dve objašnjavajuće promenljive, ocenjuje se pomoćna regresija. Ako postoji jaka linearna korelacija među ovim promenljivim, u pomoćnoj regresiji standardne greške ocena parametara biće male, a koeficijent determinacije visoke vrednosti



PRIMIJENJENA EKONOMETRIJA

Rešavanje problema

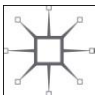
- Pojedini ekonometričari predlažu da se ovaj problem ignoriše. Naime, u svakom modelu će postojati određeni stepen multikolinearnosti, naročito u modelima vremenskih serija.



PRIMIJENJENA EKONOMETRIJA

Rešavanje problema

- Najjednostavniji načini za rešenje ovog problema su:
 - (a) izostaviti jednu od varijabli, koja izaziva problem
 - (b) Koristiti podatke koji se dobijaju transformacijom polaznih podataka (svi podaci se dele sa promenljivom koja stvara problem ili se koriste prve difference promenljivih)
 - (c) proširiti uzorak



PRIMIJENJENA EKONOMETRIJA

Primer

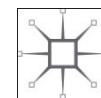
Dati su kvartalni podaci za

Uvoz (IMP)

Bruto domaći proizvod (GDP)

Indeks potrošačkih cena (CPI) i

Indeks proizvođačkih cena (PPI)

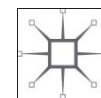


PRIMIJENJENA EKONOMETRIJA

Primer

Matrica koeficijenata korelacije

	IMP	GDP	CPI	PPI
IMP	1	0.979	0.916	0.883
GDP	0.979	1	0.910	0.899
CPI	0.916	0.910	1	0.981
PPI	0.883	0.8998	0.981	1



PRIMIJENJENA EKONOMETRIJA

Primer – samo CPI

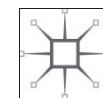
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.631870	0.344368	1.834867	0.0761
LOG(GDP)	1.926936	0.168856	11.41172	0.0000
LOG(CPI)	0.274276	0.137400	1.996179	0.0548
R-squared	0.966057	Mean dependent var	10.81363	
Adjusted R-squared	0.963867	S.D. dependent var	0.138427	
S.E. of regression	0.026313	Akaike info criterion	-4.353390	
Sum squared resid	0.021464	Schwarz criterion	-4.218711	
Log likelihood	77.00763	F-statistic	441.1430	
Durbin-Watson stat	0.475694	Prob(F-statistic)	0.000000	



PRIMIJENJENA EKONOMETRIJA

Primer –CPI sa PPI

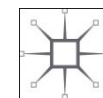
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.213906	0.358425	0.596795	0.5551
LOG(GDP)	1.969713	0.156800	12.56198	0.0000
LOG(CPI)	1.025473	0.323427	3.170645	0.0035
LOG(PPI)	-0.770644	0.305218	-2.524894	0.0171
R-squared	0.972006	Mean dependent var	10.81363	
Adjusted R-squared	0.969206	S.D. dependent var	0.138427	
S.E. of regression	0.024291	Akaike info criterion	-4.487253	
Sum squared resid	0.017702	Schwarz criterion	-4.307682	
Log likelihood	80.28331	F-statistic	347.2135	
Durbin-Watson stat	0.608648	Prob(F-statistic)	0.000000	



PRIMIJENJENA EKONOMETRIJA

Primer – samo PPI

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.685704	0.370644	1.850031	0.0739
LOG(GDP)	2.093849	0.172585	12.13228	0.0000
LOG(PPI)	0.119566	0.136062	0.878764	0.3863
R-squared	0.962625	Mean dependent var	10.81363	
Adjusted R-squared	0.960213	S.D. dependent var	0.138427	
S.E. of regression	0.027612	Akaike info criterion	-4.257071	
Sum squared resid	0.023634	Schwarz criterion	-4.122392	
Log likelihood	75.37021	F-statistic	399.2113	
Durbin-Watson stat	0.448237	Prob(F-statistic)	0.000000	



PRIMIJENJENA EKONOMETRIJA

Primer – pomoćna regresija

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.542357	0.187073	-2.899177	0.0068
LOG(CPI)	0.974766	0.074641	13.05946	0.0000
LOG(GDP)	0.055509	0.091728	0.605140	0.5495
R-squared	0.967843	Mean dependent var	4.552744	
Adjusted R-squared	0.965768	S.D. dependent var	0.077259	
S.E. of regression	0.014294	Akaike info criterion	-5.573818	
Sum squared resid	0.006334	Schwarz criterion	-5.439139	
Log likelihood	97.75490	F-statistic	466.5105	
Durbin-Watson stat	0.332711	Prob(F-statistic)	0.000000	

