

Mikroekonometrija

Ekonometrija, Master studije

Predavač: Aleksandra Nojković

Beograd, školska 2021/21

Teme I dela kursa

- 1) Metod IP i primena u mikroekonometrijskim istraživanjima

- 2) Modeli specifične zavisne promenljive:
 - a) Diskretna zavisna promenljiva (binarna ili više poređanih ishoda)
 - b) Zavisna promenljive sa ograničenjem (odsečena ili cenzurisana raspodela zavisne promenljive; modeli korekcije uzoračkog izbora).

- 4) Metode koje odgovaraju na pitanja “cause-and-effect”

Literatura

Osnovna:

- Greene, W.H. (2018), *Econometric Analysis*, 8th edition, Pearson.
- Angrist, J. and Pischke, J-S. (2014), *Mastering 'Metrics: The Path from Cause to Effects*, Princeton University Press.
- Nojković, A. (2017). *Odabrane ekonometrijske teme: metodologija i primena*, Ekonomski fakultet, Beograd.

Dopunska:

- Wooldridge, J. M. (2011), *Econometric Analysis of Cross Section and Panel Data*, 2nd edition, The MIT Press.
- Maddala, G.S. (1983), *Limited-Dependent and Qualitative Variables in Econometrics*, Cambridge University Press.

Metod instrumentalnih promenljivih

Struktura predavanja:

- **Narušena jedna od pretpostavki KLRM: slučajne objašnjavajuće promenljive.**
- Kada dolazi do narušavanja ove pretpostavke?
- Kako se to odražava na ocene parametara i na standardne greške ocena?
- Šta raditi u slučaju kada je pretpostavka o nekorelisanosti objašnjavajuće promenljive i slučajne greške narušena?

Jednostavna regresiona analiza

- *Regresiona analiza* predstavlja osnovni metodološki okvir ekonometrijskog modeliranja.
- Jednostavan model:

$$Y_i = \underbrace{\beta_0 + \beta X_i}_{\text{sistematski deo}} + \underbrace{\varepsilon_i}_{\text{stohastički deo}} \quad \text{za } i=1,2,\dots, n.$$

gde je zavisna promenljiva Y za sve opservacije i iskazana kao funkcija samo jedne nezavisne promenljive (X_i) i greške modela ε_i (često i kao u_i).

Pretpostavke KLRM

1. $E(\varepsilon_i) = 0$
2. $\text{Var}(\varepsilon_i) = \sigma^2 < \infty$
3. $\text{Cov}(\varepsilon_i, \varepsilon_j) = 0$ za i različito od j
4. **Objašnjavajuće promenljive nisu određene stohastičkim članom**
5. $\varepsilon_i \sim N(0, \sigma^2)$
6. Ne postoji tačna linearna zavisnost između objašnjavajućih promenljivih (višestruki KLRM)

Pretpostavka 4: Objasnjavajuće promenljive nisu slučajne promenljive

- Objasnjavajuće promenljive uzimaju fiksirane vrednosti iz ponovljenih uzoraka:

$$\text{Cov}(\varepsilon_i, X_i) = 0, \text{ za svako } i = 1, 2, \dots;$$

- Kada je ta pretpostavka narušena

$$\text{Cov}(\varepsilon_i, X_i) \neq 0, \text{ za svako } i.$$

tada su objasnjavajuće promenljive slučajne.

Stohastički regresori

- Slučajna objašnjavajuća promenljiva može biti:

1. Nezavisna od slučajne greške:

$$\text{cov}(X_i, \varepsilon_j) = 0, \text{ za svako } i, j$$

- Ocene su sa poželjinim svojstvima.

2. Korelisana sa slučajnom greškom za različite opservacije:

$$\text{cov}(X_i, \varepsilon_j) \neq 0, \text{ za } i \neq j$$

- Ocene su pristrasne, ali konzistentne.

3. Korelisana sa slučajnom greškom za istu jednicu posmatranja:

$$\text{cov}(X_i, \varepsilon_j) \neq 0, \text{ za } i = j$$

- Ocene su pristrasne i nekonzistentne (pokazati!).

Kada je ta pretpostavka narušena?

- **Dinamički model:**

$$y_t = b_0 + b_1 x_{1t} + \gamma_1 y_{t-1} + \varepsilon_t$$

Kako je y_t korelisano sa ε_t , to je y_{t-1} korelisano sa ε_{t-1} . Objasnjavajuća promenljiva y_{t-1} je slučajna.

- **Simultane zavisnosti:** U jednačini za y_t objašnjavajuća promenljiva x_t je slučajna (SSJ).
- **Izostavljanje relevantne promenljive.**
- Postoji greška u merenju promenljivih.

Jednostavna regresiona analiza

- *Regresiona analiza* predstavlja osnovni metodološki okvir ekonometrijskog modeliranja.
- Jednostavan model:

$$Y_i = \underbrace{\beta_0 + \beta X_i}_{\text{sistematski deo}} + \underbrace{\varepsilon_i}_{\text{stohastički deo}} \quad \text{za } i=1,2,\dots, n.$$

gde je zavisna promenljiva Y za sve opservacije i iskazana kao funkcija samo jedne nezavisne promenljive (X_i) i greške modela ε_i (cesto i kao u_i).

Metod ONK (izvođenje)

- Metod se zasniva na uslovu minimiziranju zbira kvadrata reziduala (funkcija b_0, b):

$$S(b_0, b) = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 = \sum_{i=1}^n (Y_i - b_0 - bX_i)^2.$$

Postupak:

- 1) prve parcijalne izvode po nepoznatim veličinama b_0, b izjednačiti sa nulom i
- 2) pokazati da su drugi parcijalni izvodi pozitivni.

Metod ONK (II)

- Prvi parcijalni izvodi (sistem normalnih jednačina):

$$\frac{\partial S(b_0, b)}{\partial b_0} = -2 \sum_{i=1}^n (Y_i - b_0 - bX_i) = 0 \Rightarrow \sum_{i=1}^n Y_i = nb_0 + b \sum_{i=1}^n X_i$$

$$\frac{\partial S(b_0, b)}{\partial b} = -2 \sum_{i=1}^n X_i(Y_i - b_0 - bX_i) = 0 \Rightarrow \sum_{i=1}^n X_i Y_i = b_0 \sum_{i=1}^n X_i + b \sum_{i=1}^n X_i^2$$

- Drugi parcijalni izvodi su uvek pozitivni:

$$\frac{\partial^2 S(b_0, b)}{\partial b_0^2} = 2n, \quad \frac{\partial^2 S(b_0, b)}{\partial b^2} = 2 \sum_{i=1}^n X_i^2, \quad \frac{\partial^2 S(b_0, b)}{\partial b_0 \partial b} = 2 \sum_{i=1}^n X_i.$$

Ocene metodom ONK

$$\hat{\beta}_0 = b_0 = \bar{Y} - b\bar{X}$$

$$\hat{\beta} = b = \frac{n \sum_{i=1}^n X_i Y_i - \sum_{i=1}^n X_i \sum_{i=1}^n Y_i}{n \sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i \right)^2} = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2}$$

Metodom ONK postavlja se prava za koju važi:

1)
$$\sum_{i=1}^n e_i = 0$$

2)
$$\sum_{i=1}^n X_i e_i = 0$$

Pokazati....

Metod instrumentalnih promenljivih

- Polazeći od jednostavnog KLRM:

$$Y_i = \beta_0 + \beta X_i + \varepsilon_i,$$

može se do doći do normalnih jednačina množenjem jednačine sa 1 i sumiranjem, odnosno sa X_i i sumiranjem.

- 1) $\sum Y_i = n\beta_0 + \beta \sum X_i + \sum \varepsilon_i$
- 2) $\sum X_i Y_i = \beta_0 \sum X_i + \beta \sum X_i^2 + \sum \varepsilon_i X_i.$

- Očekivana vrednost ovih jednačina: zamena parametara odgovarajućim ocenama i izjednačavanjem poslednjih sabiraka sa nulom, dobijamo uobičajeni sistem normalnih jednačina.
- 1) jednačina je u redu, ali druga nije tačna ukoliko $\sum \varepsilon_i X_i \neq 0.$

Metod instrumentalnih promenljivih

- Sistem normalnih jednačina postaje:

$$\sum_{i=1}^n Y_i = n\dot{b}_0 + \dot{b} \sum_{i=1}^n X_i$$

$$\sum_{i=1}^n Z_i Y_i = \dot{b}_0 \sum_{i=1}^n Z_i + \dot{b} \sum_{i=1}^n Z_i X_i,$$

gde je Z_i promenljiva koja je istovremeno nekorelisana sa ε_i i visoko korelisana sa X_i

- Ocena metodom IP:

$$\dot{b} = b_{iv} = \frac{\sum_{i=1}^n Z_i Y_i}{\sum_{i=1}^n Z_i X_i}$$

Metod instrumentalnih promenljivih (II)

- Instrument Z_i možemo tretirati kao način da „izdvojimo“ onaj deo X_i koji nije korelisan sa ϵ_i .

$$\sum_{i=1}^n Y_i = n\hat{b}_0 + \hat{b} \sum_{i=1}^n X_i$$

$$\sum_{i=1}^n Z_i Y_i = \hat{b}_0 \sum_{i=1}^n Z_i + \hat{b} \sum_{i=1}^n Z_i X_i,$$

- U matričnoj notaciji, vektor ocena metodom IV (IP):

$$b_{IV} = (Z'X)^{-1}Z'Y,$$

gde je Z vektor instrumentalnih promenljivih, istih dimenzija kao matrica X (egzogeni regresori su sami sebi instrumenti).

Ocene metodom IP (IV)

- Ocene dobijene metodom IP su **pristrasne i konzistentne**.

$$\hat{b} = \frac{\sum_{i=1}^n z_i y_i}{\sum_{i=1}^n z_i x_i} = \frac{\sum_{i=1}^n z_i Y_i}{\sum_{i=1}^n z_i x_i} = \frac{\sum_{i=1}^n z_i (\beta_0 + \beta X_i + \varepsilon_i)}{\sum_{i=1}^n z_i x_i} = \beta + \frac{\sum_{i=1}^n z_i \varepsilon_i}{\sum_{i=1}^n z_i x_i}$$

$$E(\hat{b}) = \beta + E \left[\frac{1}{\sum_{i=1}^n z_i x_i} (z_1 \varepsilon_1 + z_2 \varepsilon_2 + \dots + z_n \varepsilon_n) \right] \neq \beta.$$

$$\hat{b} = \beta + \frac{\sum_{i=1}^n z_i \varepsilon_i / n}{\sum_{i=1}^n z_i x_i / n}$$

$$E(\hat{b}) = \beta + E \left(\frac{\sum_{i=1}^n z_i \varepsilon_i / n}{\sum_{i=1}^n z_i x_i / n} \right) = \beta + \frac{0}{\text{konstanta}} = \beta, n \rightarrow \infty.$$

- Slično se može pokazati i za ocenu $\hat{b}_0 = \bar{Y} - \hat{b} \bar{X}$.

Metode za prevazilaženje ovog problema

- Ocene dobijene metodom ONK su pristrasne i nekonzistentne (pokazati...)
- Bez obzira na uzrok koji je doveo do korelnosti objaš. prom. i slučajne greške, problem se prevazilazi metodom instrumentalnih promenljivih (podsecanje: metod ocenjivanja SSJ).
- Ocene dobijene metodom IP (IV) su pristrasne, ali konzistentne (pokazati...)

Endogenost i egzogenost: pojam u kontekstu metoda IP

- Metod IV razlikuje dve vrste promenljivih:
 - 1) Promenljive korelisane sa slučajnom greškom – **endogene**.
 - 2) Promenljive koje nisu korelisane sa slučajnom greškom modela – **egzogene**.
- Istorijski, ovi termini potiču iz modela SSJ gde se endogenim promenljivima nazivaju one determinisane unutar sistema, dok su egzogene određene spolja (videte više u delu o SSJ).

Izbor instrumenata

1. Instrumentalna promenljiva je prethodna vrednost objašnjavajuće promenljive X_{i-1} , X_{i-2}, \dots (ukoliko ne postoji autokorelacija).
2. Metod dve grupe: podaci objašnjavajuće promenljive se urede rastućim vrednostima i podele na dve grupe-iznad i ispod medijane (vrednosti -1 i 1).
3. Metod tri grupe: isto, ali se centralnim podacima dobijenim metodom dve grupe dodeli vrednost 0.
4. Instrumentalna promenljiva uzima vrednosti rednog broja opservacija objašnjavajuće promenljive uređenih po veličini ($i=1, 2, \dots, n$).
5. **Instrumenti – promenljive van modela!**

Testovi relevantni pri primeni metoda IV/IP

- Da li je data objašnjavajuća promenljiva zaista **endogena**?

$$\text{cov}(X_i, \varepsilon_i) \neq 0.$$

- Hausman-ov test.

- Ako postoji problem endogenosti, kako odabrati skup instrumentalnih promenljivih koji je **validan**?

- Testovi koji razmatraju dve teme.

Napomena: Svi testovi definisani u nastavku podrazumevaju **homoskedastične** i **neautokorelisane** slučajne greške.

Validnost izabranih instrumenata

- Metod IP daje pouzdane ocene parametara ukoliko instrumenti ispunjavaju sledeće:

- 1) Nisu korelisani sa slučajnom greškom (**egzogene priroda instrumenata**):

$$\text{cov}(Z_i, \varepsilon_i)=0.$$

- Hansen-Sarganov J test.

- 2) Korelisani su sa objašnjavajućom endogenom promenljivom (instrumenti su **relevantni, nisu slabi**):

$$\text{cov}(Z_i, X_i)\neq 0.$$

- Stock-Watson –ov i Stock-Yogo-ov test (testovi za proveru loših / slabih instrumenata).

Svojstva ocena dobijenih različitim metodama ocenjivanja

Metod	Egzogenost	Endogenost
ONK	konzistentna, efikasna	nekonzistentna
IP/IV	konzistentna, neefikasna	konzistentna

Provera endogenosti: Hausman-ov test

- Definišemo hipoteze:

$H_0: \text{cov}(X_i, \varepsilon_i) = 0$, X_i je egzogeno.

$H_1: \text{cov}(X_i, \varepsilon_i) \neq 0$, X_i je endogeno.

- Formalno tvrđenje H_0 je da među ocenjenim koeficijentima nema sistematske razlike:

$$d = \hat{\beta}_{IP} - \hat{\beta}_{ONK} = 0$$

- Testiranje se zasniva na Wald-ovoj test statistici:

$$H = d' [Asim.Var(d)]^{-1} d,$$

koja ima χ^2 raspodelu sa brojem stepeni slobode koji je jednak broju promenljivih čiju endogenost ispitujemo.

Napomena: Pretpostavlja se da je **skup instrumentalnih** promenljivih **adekvatno određen!**

Hausman-ov test specifikacije (nastavak)

- Za izračunavanje asim. varijanse Hausman je predložio jednostavan postupak (usled nezavisnosti ovako dobijenih ocena):

$$Asim. Var(d) = Asim. Var\left(\hat{\alpha}_{IP}\right) - Asim. Var\left(\hat{\alpha}_{ONK}\right).$$

- Postoji jednostavnije verzija ovog testa, poznatija kao **Wu-ov test dodatih regresora** – proverava sa endogenost pojedinačnih objašnjavajućih promenljivih (detaljnije u nastavku).

Druga verzija Hausman-ovog testa (Wu-ov test)

- Posmatramo jednačinu:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i.$$

- Neka su Z_{1i} i Z_{2i} instrumentalne promenljive (npr. pomaknute vrednosti zavisne i objašnjavajuće promenljive).
- Ispitujemo endogenost objašnjavajuće promenljive X_i , odnosno proveravamo da li je X_i korelisano sa ε_i .

Druga verzija Hausman-ovog testa (II)

1. Ocenjujemo promenljivu X_i kao funkciju instrumenata primenom metoda ONK:

$$X_i = a_0 + a_1 Z_{1i} + a_2 Z_{2i} + v_i$$

i dobijamo rezidualne u_i .

2. Polaznu jednačinu proširujemo ovako dobijenim rezidualima (ili ocenom za X_i iz prvog stepena) i potom takav model ocenjujemo primenom metoda ONK:

$$Y_i = \beta_0 + \beta_1 X_i + \delta u_i + \varepsilon_i.$$

$H_0 : \delta = 0$, Promenljiva X_i nema endogeni karakter

$H_1 : \delta \neq 0$, Promenljiva X_i ima endogeni karakter

3. Testiramo validnost datih hipoteza primenom t ili F statistike.

Hausman-ov test (nastavak)

- Ukoliko se proverava endogeni status više promenljivih, onda se jednačina proširuje sa rezidualima svih korespondirajućih jednačina ocenjenih u prvom stepenu (sve objašnjavajuće promenljive čiju endogenost proveravamo u prvom stepenu se ocene kao funkcije instrumenata i sačuvaju reziduali).
- Njihova zbirna signifikantnost se ispituje primenom F-testa.
- Reč je o modifikovanoj verziji testa (Durbin-Wu-Hausman, DWH), pri čemu EViews izračunava vrednost statistike na bazi nešto drugačije aproksimacije.

Hausman-ov test (nastavak)

- U obe verzije testa hipoteze definišemo kao:

H_0 : Ocene dobijene po oba metoda (ONK i IP) su konzistentne, pri čemu su ocene ONK relativno efikasnije (razmatrana promenljiva je egzogena).

H_1 : Ocene dobijene metodom IP su konzistentne, dok metod ONK daje nekonzistentne ocene (razmatrana promenljiva je endogena).

Provera validnosti instrumenata

- Primena metoda IP (2SNK – **objasniti!!!**) daje pouzdane ocene parametara ukoliko su izabrani instrumenti:
 - 1) Relevantni (korelisani sa endogenom promenljivom čiji su instrumenti – ocena u prvom stepenu).
 - 2) Egzogeni (nisu korelisani sa slučajnom greškom polazne jednačine).

Provera validnosti izabranih instrumenata

- Posmatramo jednačinu:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i.$$

- X_i je objašnjavajuća endogena promenljiva.
- Neka su Z_{1i} i Z_{2i} potencijalne instrumentalne promenljive koje razmatramo. Postavljaju se dva pitanja:

- 1) Da li su Z_{1i} i Z_{2i} relevantne instrumentalne promenljive?
- 2) Da li su Z_{1i} i Z_{2i} egzogene promenljive?

Test relevantnosti: Stock-Watson-ov pristup

1. U prvom stepenu ocenimo endogenu objašnjavajuću promenljivu X_i kao funkciju instrumenata:

$$X_i = a_0 + a_1 Z_{1i} + a_2 Z_{2i} + v_i$$

2. Proveravamo značajnost zajedničkog uticaja objašnjavajućih promenljivih (instrumenata) primenom F-testa ($H_0: a_1=a_2=0$).

3. Ako je vrednost F-statistike veća od 10 smatramo da su instrumentalne promenljive relevantne (ako su ipak statistički značajne, ali F-statistika manja od 10 – **ipak su instrumenti slabi!!!**).

4. **Pravilo „palca“** – F-statistika veća od 10, tada je pristrasnost oko 10% pristrasnosti ocene ONK (zanemarljivo) – u slučaju **jedne endogene promenljive!**

Test relevantnosti: Stock-Yogov pristup

- Bazira se na primeni Cragg-Donald-ovog testa, pri čemu se ostvarena vrednost poredi sa skup kritičnih vrednosti koje su predložili **Stock i Yogo**.
- Navode se odgovarajuće kritične vrednosti (uobičajeno navode za 10, 15, 20 i 25%), a zavise od broja endogenih promenljivih i broja instrumenata u razmatranoj jednačini.
- Manja ostvarena vrednost od kritične vrednosti sugerije prihvatanje nulte hipoteze da su izabrani instrumenti loši.

Test egzogenosti instrumenata: J-test

1) Ocenjujemo polaznu jednačinu:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i,$$

i sačuvamo rezidualne e_i .

2) Primenom metoda ONK ocenjujemo sledeću pomoćnu regresiju iz koje dobijamo vrednost R^2 :

$$e_i = a_0 + a_1 Z_{1i} + a_2 Z_{2i} + v_i.$$

3) Proveravamo validnost hipoteza:

H_0 : $a_1 = a_2 = 0$ (instrumentalne promenljive su egzogene).

H_1 : H_0 nije tačna (instrumentalne promenljive nisu egzogene).

Test egzogenosti instrumenata: J-test (nastavak)

4) Formiramo sledeću statistiku: $J = TR^2$,

za koju se može pokazati da poseduje χ^2 raspodelu sa brojem stepeni slobode koji je razlika između broja instrumentalnih promenljivih (uključujući konstantu) i broja parametara polazne jednačine.

5) Ako je izračunata vrednost J test-statistike veća od odgovarajuće kritične vrednosti odbacuje se H_0 o egzogenom karakteru instrumentalnih promenljivih (instrumenti **nisu korektno odabrani**).

Napomena: Test se koristi samo kod **prekomerne identifikovanosti** (veći broj instrumenata od broja objašnjavajućih prome. polazne jednačine) i može se smatrati testom za proveru ovakve identifikovanosti (više u delu o SSJ).

Primer 1: Izbor instrumenata

- U udžbeniku Wooldridge (2012) korišćeni su podaci o zaradama zaposlenih, udatih žena:

$$\log(wage) = \beta_0 + \beta_1 educ + u$$

- Sposobniji pojedinci su uspješniji na tržištu rada i pri tome duže se školuju. Potvrđena je *endogenost* objašnjavajuće prom. *educ* (H – stat = 1.67), odnosno *educ*je korelisano sa izostavljenom prom. tzv. „urođenim“ sposobnostima.
- Radi kasnijeg poređenja date su ocene ONK (n=428):

$$\log(\widehat{wage}) = -0.185 + 0.109 educ; R^2 = 0.118$$

(0.185) (0.014)

- Efekat dodatne godine školovanja je ocenjen oko 11%.

Primer 1: Izbor instrumenata (nastavak)

- U nastavku je umesto nivoa obrazovanja (*educ*), koji su korelisani sa slučajnom greškom, kao instrument korišćena njihova ocena u funkciji nivoa obrazovanja roditelja (nije korelisano sa greškom, a jeste sa *educ*).
- Ocene dobijene metodom IP:

$$\log(\widehat{wage}) = \dots + 0.061 \text{ educ}; R^2 = 0.039$$

(0.031)

- Efekat dodatne godine školovanja je ocenjen kao dvostruko manji (6.1%), ali širi interval poverenja koji obuhvata i ocene ONK (nismo sig. da je ova razlika i statistički značajna). Sa tri instrumenta (uključeno je i obrazovanje supružnika) ovaj efekat je ocenjen oko 8%.

Primer 2: Izbor instrumenata

- Angrist and Krueger, „Does Compulsory School Affect Schooling and Earnings“, *Quarterly Journal of Economics*, 1991.
- Rad odgovara na pitanje da li obavezno školovanje utiče na odluku da se završi srednja škola, odnosno da li utiče na zarade.
- Rad koji se često navodi u **analizi izbora instrumentalnih promenljivih i provere validnosti** izabranih instrumenata (detaljnije u nastavku).

Izbor instrumenata (nastavak)

Ocenjuje se logaritam zarada kao funkcija godina školovanja i seta drugih objašnjavajućih promenljivih.

- Greška/reziduali te jednačine korelisani sa promenljivom godine školovanja?
- U tom slučaju ocene metodom ONK (OLS) su pristrasne i nekonzistentne (poznato kao problem izostavljanja relevantne promenljive)
- Koja prom. je korelisana sa brojem godina školovanja a nije sa greškom u jednačini zarada – **datum rođenja!**

Model ocenjen metodom 2SNK

- AK(1991) na bazi podataka iz popisa u SAD-u, st. 997 ocenjuju sledeći sistem:

$$(1) \quad E_i = X_i \pi + \sum_c Y_{ic} \delta_c + \sum_c \sum_j Y_{ic} Q_{ij} \theta_{jc} + \epsilon_i$$

$$(2) \quad \ln W_i = X_i \beta + \sum_c Y_{ic} \xi_c + \rho E_i + \mu_i,$$

pri čemu E_i odnosi na nivo obrazovanja i -tog pojedinca, X_i je vektor objašnjavajućih promenljivih, Q_{ij} je veštačka promenljiva koja se odnosi na kvartal rođenja ($j=1,2,3$), dok je Y_{ic} je veštačka koja se odnosi na godinu rođenja ($c=1,2,\dots,10$), a W_i je nedeljna zarada.

- Ukoliko je broj godina školovanja u ovoj jednačini korelisan sa slučajnom greškom (μ_i), ocene ONK su pristrasne (pristrasnost koja nastaje usled endogenosti, odnosno usled *izostavljanja relevantne promenljive*).

Izbor instrumenata (objašnjenje)

- Pri tome, obavezno školovanje je do navršene 16 godine, tako ukoliko osoba napuni 16 godina u januaru, ona može napustiti školu pre završetka poslednje godine.
- Ukoliko pak navršava 16 godina u julu, završiće poslednju godinu pre nego što uopšte bude u mogućnosti da školu napusti.
- Kako nije logično da na nivo zarade utiče period godine u kojoj je osoba rođenja, to se rešava ocenjivanjem u dva stepena (zamenom promenljive godine školovanja odgovarajućim instrumentom, odnosno ocenom iz prvog stepena).

Izbor instrumenata (nastavak)

- Iz tog razloga, promjenljiva godine školovanja se ocenjuje u prvom stepenu kao funkcija seta objašnjavajućih promjenljivih i veštačkih promjenljivih, koje se odnose na kvartal i godinu rođenja pojedinca (jednačine 1 i 2, na strani 997 u orig. radu).
- **Ocena** godina školovanja **iz prvog stepena** se koristi kao instrument u drugom stepenu (stvarne godine školovanja se zamenjuju ovom ocenom godina školovanja iz prvog stepena i ponovo se primenjuju ONK).
- Pretpostavlja se da navedeno predstavlja dobar instrument, jer je korelisano sa stvarnim godinama školovanja a nije sa greškom u jednačini zarada (tako da jednačina **nema problem endogenosti**).
- Naravno, ostaje **pitanje kvaliteta ovih instrumenata za ocenjivanje broja godina školovanja**.

Kritika izbora instrumenata

- Ne postoje dokazi za odbacivanje hipoteze o validnosti izabranog seta instrumenata (instrumenti su egzogeni).
- **Dakle, skup instrumentalnih promenljivih je adekvatno određen (različito od kasnije razmatranog pitanja jačine (kvaliteta) izabranih instrumenata)!!!**
- U renomiranom udžbeniku autora Stock i Watson (*Introduction to Econometrics, 2003*), upravo je rad *Angrist i Krueger (1991)*, naveden kao primer tzv. „Scary Regression“, odnosno ilustracija za primer slabih/loših instrumenata.

Kritika izbora instrumenata (nastavak)

- Ovo je dokazano na veoma originalan način, tako što je ispitanicima na slučaj dodeljen kvartal (sezona) rođenja i sprovedeno ponovno ocenjivanje metodom IP/2SNK.
- Pokazalo se da je primena metoda 2SNK **dala iste rezultate kao kada su korišćeni „slučajni“ datumi rođenja pojedinaca.**
- **Pogledati:** Kritika rada i ponovljena analiza na istim podacima korišćenjem „*fake instruments*“ u radu *Bound, Jaeger and Baker* (1995).
- U radu AK, vrednost F-statistike u nekim ocenjenim jednačinama prvog stepena je manji od 2 (engl. *rule-of-thumb* zahteva vrednost iznad 10). Ocene pristrasne, moguće i više od ONK za male uzorke!