

Heteroskedastiivsus Jääkide autokorrelatsioon Endogeensus ja instrumenttunnused

Brooks, Introductory Econometrics for Finance, Ch. 5.1-5.7, 13.4.

Wooldridge, J. M., Introductory Econometrics, 6th ed., Ch. 8.4 Weighted Least Squares Estimation.

Wang, P. Instrumental Variables Approach to Correct for Endogeneity in Finance. In Lee, C.-F., Lee, J. (eds.), Handbook of Financial Econometrics and Statistics. Ch 95.

Loengu teemad

- Heteroskedastiivsus
 - Heteroskedastiivsuse mõiste, mõju, testimine
 - Bootstrap meetodil leitud standardvead
 - Kaalutud vähimruutude meetod
- Jääkliikmete autokorrelatsioon
 - Autokorrelatsiooni mõju, testimine
 - Dünaamilised mudelid
- Endogeensus ja instrumenttunnused
 - Regressorite eksogeensus
 - Instrumendid
 - Kahe-etapiline vähimruutude meetod

OLS klassikalised Gauss-Markov eeldused

1. Juhuslike liikmete keskväärtus $E[u_i] = 0$

2. Regressorid peavad olema eksogeensed, $\text{Cov}(u_i, X_{aj}) = 0$

3. Homoskedastiivsus: $\text{Var}(u_i) = \text{const}$

4. Juhuslike liikmete autokorrelatsiooni puudumine: $\text{Cov}(u_t, u_{t+j}) = 0$

Eelduse 1 täitmise garanteerib vabaliikme esinemine mudelis.

Kui eeldus 2 pole täidetud, on parameetrite hinnangud nihkega.

Kui eeldused 3 või 4 pole täidetud, on parameetrite standardvigade hinnangud nihkega.

Mõningaid heteroskedastiivsuse põhjuseid

- Matemaatilise mudeli vale kuju.
 - Näiteks log-log mudeli asemel hinnatakse lineaarset mudelit.
- Mõni oluline seletav tunnus on mudelist välja jäänud.
- Üks või mitu seletavat tunnust on asümmeetrilised.
- Üksikute erindite (*outliers*) esinemine vaatluste hulgas.
- Andmekogumismeetodid paranevad -> vaatlusvigade suurus väheneb, st juhuslikud liikmed vähenevad.
- Muud põhjused:
 - suurema kasumiga ettevõtetel on dividendipoliitikas suuremad erinevused;
 - suurema sissetulekuga peredel on säästmisharjumused rohkem hajunud.

HETEROSKEDASTIIVSUST VÕIB PÕHJUSTADA NII MUDEL KUI KA ANDMED

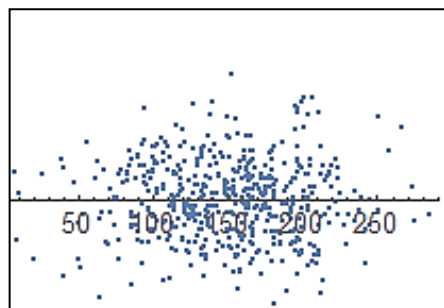
Heteroskedastiivsuse mõju

- Vealiikmete dispersioon EI ESINE parameetrite HINNANGUTE arvutusvalemities.
- Järelikult heteroskedastiivsus parameetrite hinnanguid ei mõjuta. Need on ikka **nihketa**.
- Vealiikmete dispersioon ESINEB parameetrite hinnangute STANDARDVIGADE arvutusvalemities.
 - Parameetrite hinnangud **ei ole enam efektiivsed**.
 - Parameetrite usalduspiirid tulevad valed.
 - Mudeli ja parameetrite olulisuse testimine (F -test ja t -testid) võivad anda valesid tulemusi.
- Järelikult võime teha valesid järeldusi mudelisse kuuluvate või mittekuuluvate tunnuste suhtes.

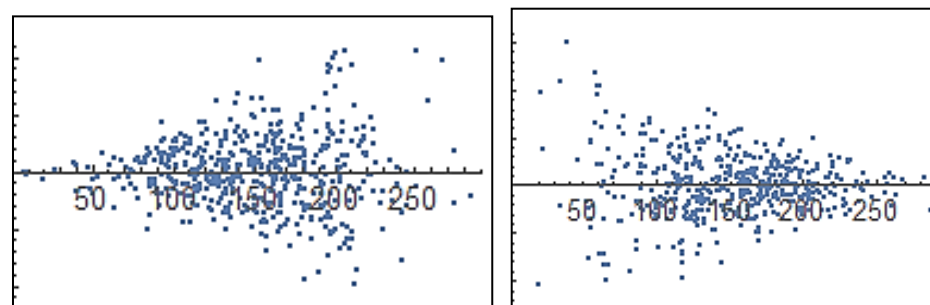
Heteroskedastiivsuse testimine

- Jääkide graafiline analüüs
 - y-teljel mudeli jääkliikmed, nende ruudud või absoluutväärtused
 - x-teljel mõni valitud sõltumatu muutuja X_i või siis sõltuva muutuja hinnang Y_i

Homoskedastiivsus



Heteroskedastiivsus



- Formaalsed testid
 - Breusch-Pagan'i test
 - White'i test
 - ARCH test (ainult aegridade korral)
 - jt
- Formaalsete kriteeriumite kasutamisse peab suhtuma ettevaatlikult. Soovitav kasutada erinevaid testimismeetodeid ja täiendada formaalseid kriteeriume graafilise analüüsiga.

ARCH test

Saab kasutada, kui regressioonimudel on aegridade vahel.

Regressiooni jääkliikmete jaoks viiakse läbi abiregressiooni hindamine, mis on jääkliikmete ARCH(q) mudel:

$$\hat{u}_t^2 = \gamma_0 + \sum_{i=1}^q \gamma_i \hat{u}_{t-i}^2 + v_t$$

Arvestatud, et
 $\hat{\sigma}_t^2(u_t) = \hat{u}_t^2$

Nullhüpotees: jääkliikmetes ARCH-i ei esine järguni q .

$$\gamma_1 = \gamma_2 = \dots = \gamma_q = 0$$

Sellisel juhul on jääkliikmete dispersioon konstantne $\hat{\sigma}_t^2(u_t) = \hat{u}_t^2 = \gamma_0$

Teststatistik $LM = nR^2 \sim \chi^2(q)$ allub asümptootiliselt χ^2 jaotusele.

Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50(4), 987–1007.

Mida teha, kui heteroskedastiivsus esineb?

- Kontrollida mudeli spetsifikatsiooni:
 - kas mudelil on õige kuju;
 - kas mõni oluline tunnus on välja jäetud;
 - mudelit teisendada, uuesti hinnata.
- Leida heteroskedastiivsuse suhtes kohandatud standardhälvete hinnangud (White'i hinnangud, *heteroskedasticity-consistent standard errors, robust errors*)
 - Tihti esitatakse publikatsioonides mõlemad standardvigade hinnangud: tavalised ja kohandatud.
 - Need ei kaota heteroskedastiivsust, need arvestavad sellega.
- Kasutada standardvigade hindamiseks *bootstrap* meetodit.
- Kasutada hariliku vähimruutude meetodi asemel muud meetodit:
 - kaalutud vähimruutude meetod WLS (eemaldab heteroskedastiivsuse);
 - suurima tõepära meetod MLE (lubab heteroskedastiivsust, standardvigade hinnangud sellest ei sõltu).

BOOTSTRAP

Bootstrap

- *Bootstrap* meetodit kasutatakse testimiseks, dispersiooni ja usalduspiiride hindamiseks.
- Tehakse taasvalikuid (*resampling*) olemasolevast valimist.
 - Võetakse k korda m tagasipanekuga vaatlust ja arvutatakse nendest analüüsitava parameetri k hinnangut (*bootstrap replicates*).
 - Alamvalimid võivad olla sama mahuga, mis esialgne, $m=n$.
- Hinnatavate statistikute jaotus saadakse taasvalikute põhjal.
- Kogumi jaoks pole vaja mingeid eeldusi.
- Ainuke eeldus: olemasolev valim on esinduslik, esindab hästi uuritavat kogumit.
- Meetodi nimetus: parun Münchausen "*Pulling oneself up by one's bootstraps*"



Formaalne eeskiri

On valim $\{x_1, x_2, \dots, x_n\}$



1. Teeme alamvalimi $B_1 = \{x_1^*, x_2^*, \dots, x_n^*\}$

Iga x_i^* on tagasipanekuga juhuvaljavõtt. Tõenäosus $P(x = x_i^*) = \frac{1}{n}$

2. Alamvalimi B_1 põhjal leiame meid huvitava statistiku hinnangu $\hat{\theta}_1^B$

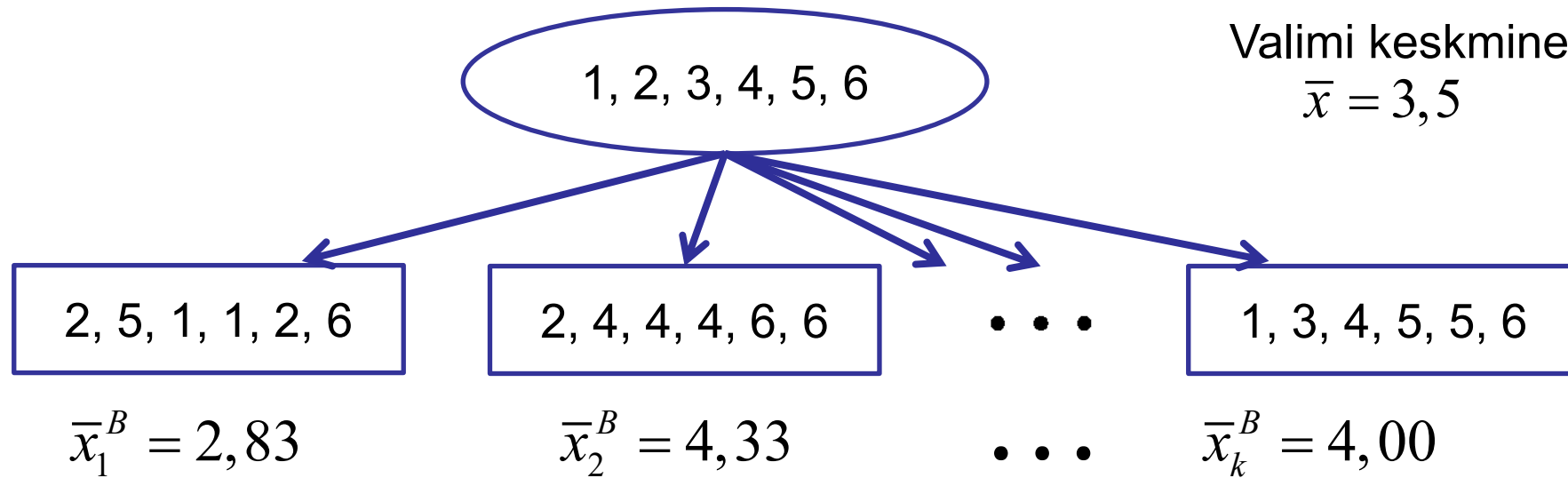
3. Kordame protseduure 1 ja 2 k korda ja saame $\{\hat{\theta}_1^B, \hat{\theta}_2^B, \dots, \hat{\theta}_k^B\}$

4. Statistiku θ jaotuseks võtame suuruste $\{\hat{\theta}_1^B, \hat{\theta}_2^B, \dots, \hat{\theta}_k^B\}$ empiirilise jaotuse.

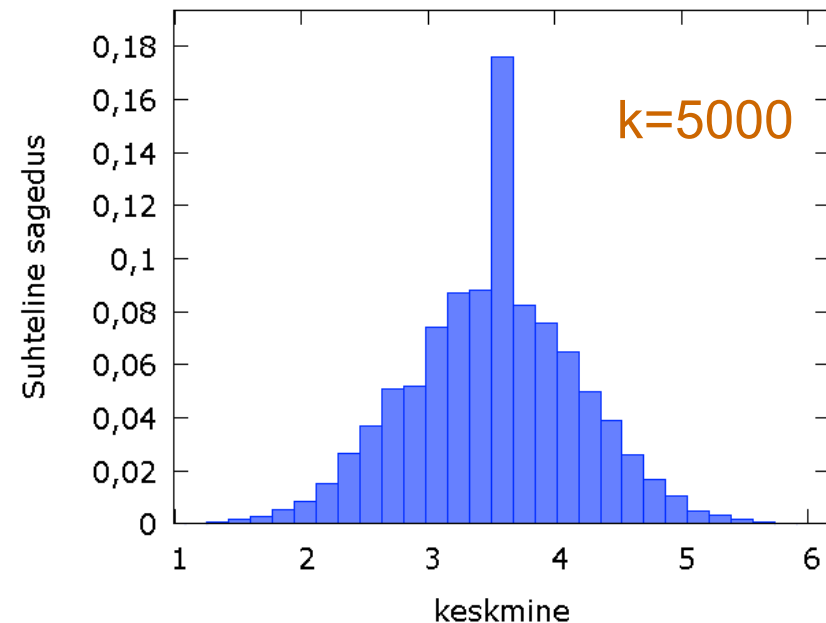
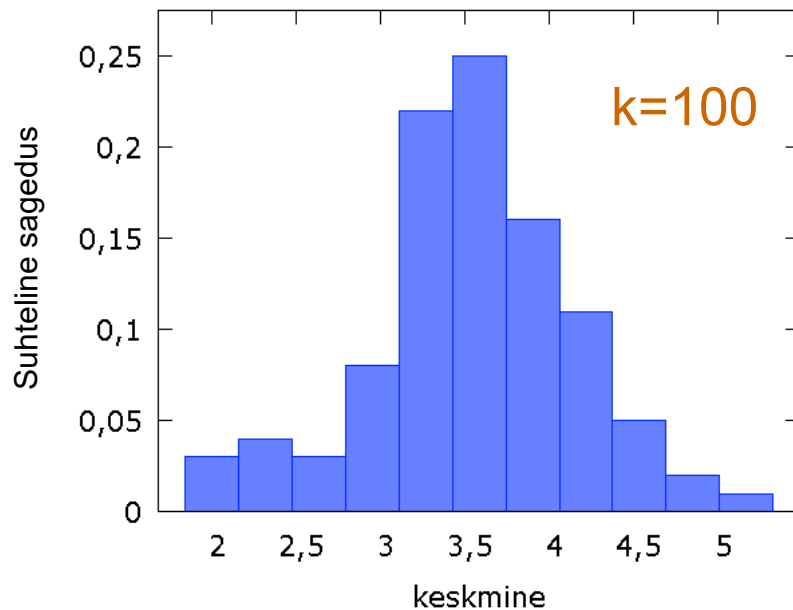
Taasvalikute teostamine

1, 2, 3, 4, 5, 6

NÄIDE: *bootstrap* meetod keskväärtuse jaotuse jaoks



Valimite keskmiste jaotus



Hinnangu standardviga *bootstrap* meetodil

1. Viiakse läbi k alamvalimi moodustamine *bootstrap* meetodil.
2. Leitakse iga valimi põhjal vajaliku parameetri (näiteks regressioonmudeli parameetri) hinnang. Saadakse hinnangute hulk

$$\{\hat{\theta}_1^B, \hat{\theta}_2^B, \dots, \hat{\theta}_k^B\}$$

3. Leitakse selle valimi standardhälve

$$se(\hat{\theta}^B) = \sqrt{\frac{1}{k-1} \sum_{i=1}^k (\hat{\theta}_i^B - \bar{\hat{\theta}}^B)^2}$$

kus
$$\bar{\hat{\theta}}^B = \frac{1}{k} \sum_{i=1}^k \hat{\theta}_i^B$$

Nii saadakse standardvigade se hinnangud kõigi parameetrite jaoks.

NÄIDE: standardvead bootstrap meetodil

20 Saksamaa ettevõtet

Kasumi sõltuvus käibest

$$profits_i = 116 + 7,89sales_i + \varepsilon_i, \quad R^2 = 0,403 \quad n = 20$$

Tavalised
standardvead

profits	Coefficient	Std. err.	t	P> t
sales	7.892176	2.263828	3.49	0.003
_cons	116.1593	50.85614	2.28	0.035

Esineb
heteroskedastiivsus

Kohandatud
standardvead

profits	Coefficient	Robust std. err.	t	P> t
sales	7.892176	4.097213	1.93	0.070
_cons	116.1593	50.6384	2.29	0.034

Ei ole statistiliselt
oluline nivool 0,05.

Standardvead
bootstrap meetodil.
Korduste arv 50.

profits	Observed coefficient	Bootstrap std. err.	z	P> z
sales	7.892176	3.973941	1.99	0.047
_cons	116.1593	57.47361	2.02	0.043

On suurem kui
tavaline, aga veidi
väiksem kui
kohandatud.

Taasvaliku valim maht

- Enamasti on valimi mahuks originaalvalimi maht n .
- See võimaldab saada antud valimi mahule n vastavad statistikud.
- Taasvaliku valimi mahtu võib muuta, kui soovime uurida näiteks testi võimsuse sõltuvust valimi mahust.

Milline on piisav korduste arv?

- Konkreetset vastust pole.
- Teoreetiliselt lõpmatu.
- Praktikas piisab sellisest, et sama korduste arvuga k seeriate korral tuleb statistiku väärtus ligikaudu sama.

Soovitus

1. Valida korduste arv (näiteks 250) ja leida *bootstrap* hinnangud.
2. Valida uus juhuarvude generaatori lähteväärtus (*seed*) ja korrata *bootstrap* hinnangute leidmist sama korduste arvuga.
 - Kui hinnangud ei erine, siis korduste arvust piisab.
 - Kui erinevad, tuleb korduste arvu suurendada.

Märkusi

- Ei ole vaja mingeid eeldusi analüüsitava statistiku jaotuse kohta.
- Millal *bootstrap* meetod ei sobi?
 - Kui valimis esineb suure omapäraga vaatlusi.
 - Kui vaatlusandmete korral esineb autokorrelatsioon.
 - Siis kasutatakse meetodit *moving block bootstrap*. Stata-s see võimalus puudub.

Vt näiteks Davidson, A.C. Hinkley D.V. (1997) Bootstrap Methods and their Application.

KAALUTUD VÄHIMRUUTUDE MEETOD

Kaalutud vähimruutude meetod WLS

- Kui **me teame**, kuidas vealiikmete dispersioon muutub, kasutame kaalutud vähimruutude meetodit (*weighted least squares*) WLS.
- Heteroskedastiivsuse esinemisel
 - WLS hinnangud on **efektiivsemad** kui OLS hinnangud;
 - t - ja F -statistikud alluvad vastavalt t - ja F -jaotusele.
- WLS-i kasutatakse ka muudel eesmärkidel, mitte ainult heteroskedastiivsuse kõrvaldamiseks.
 - Parameetrite hinnangute täpsus.
 - On võimalik saavutada täpsemad hinnangud, väiksemad standardvead ja kitsamad usaldusintervallid.
 - Prognoosimise täpsus.
 - Omistades vaatlustele erinevaid kaale, on võimalik nõ „karistada“ vähemväärtuslikke vaatlusi ja tõsta olulist informatsiooni pakkuvate vaatluste kaalu.

Vaatluste kaalumine

Esialgne mudel (OLS) $y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + \dots + u_i$

Olgu vealiikmete dispersioon $Var[u_i | X] = \sigma^2 h_i^2$ h astendaja on 2 seepärast, et peab olema positiivne.

kus h_i^2 on teada. Seega me teame, kuidas vealiikmete dispersioon muutub vaatlustest i sõltuvalt.

$\sigma^2 = const$ Selle suurust ei tea, kuid seda saab hinnata.

Korrutame mudelis iga vaatluse väärtused läbi selle vaatluse kaaluga $\frac{1}{h_i}$

Uus mudel, kus kaalutud vaatlused $\frac{y_i}{h_i} = \beta_1 \frac{1}{h_i} + \beta_2 \frac{x_{2i}}{h_i} + \beta_3 \frac{x_{3i}}{h_i} + \dots + \frac{u_i}{h_i}$

Need vaatlused, mis põhjustavad suuremat dispersiooni, on väiksema kaaluga.

Uutes tähistustes

WLS $y_i^* = \beta_1 x_{1i}^* + \beta_2 x_{2i}^* + \beta_3 x_{3i}^* + \dots + u_i^*$ $y_i^* = \frac{y_i}{h_i}, \quad x_{1i}^* = \frac{1}{h_i}, \quad x_{2i}^* = \frac{x_{2i}}{h_i}, \dots, \quad u_i^* = \frac{u_i}{h_i}$

Parameetrite hinnangud tulevad teised. $\hat{\beta}_{WLS} \neq \hat{\beta}_{OLS}$

Kaalutud mudeli vealiikmete dispersioon

WLS $y_i^* = \beta_1 x_{1i}^* + \beta_2 x_{2i}^* + \beta_3 x_{3i}^* + \dots + u_i^*$ $u_i^* = \frac{u_i}{h_i}$

$$Var[u_i^*] = Var\left[\frac{u_i}{h_i}\right] = \frac{1}{h_i^2} Var[u_i]$$

Kuna $Var[u_i] = \sigma^2 h_i^2$ kus $\sigma^2 = const$

$$Var[u_i^*] = \frac{1}{h_i^2} \sigma^2 h_i^2 = \sigma^2 = const$$

Kaalutud mudeli korral on vealiikmete u_i^* dispersioon konstantne.
Heteroskedastiivsus puudub!

Konstantse σ^2 hinnang $\hat{\sigma}^2 = \frac{1}{n-K} \sum_{i=1}^n h_i^{-2} (y_i - x_i' \hat{\beta}_{WLS})^2$

Milline on kaal?

- Praktikas reeglina juhuslike vigade dispersioone teada pole.
- Eeldatakse, et dispersioon sõltub mõnest alljärgnevast suurusest:
 - mõni seletav tunnus x_j või funktsioon sellest $f(x_j)$;
 - sõltuv tunnus y või funktsioon sellest $f(y)$;
 - mingi mudelis mitteesinev tunnus z , mille kohta on ka väärtused olemas.

Näiteks olgu meil teada, et $\sigma^2(u_i) \approx cx_i$

Kaal, millega dispersiooni kaaluda, on $\frac{1}{x_i}$

Kaal, millega juhuslikke vigu kaaluda, on $\frac{1}{\sqrt{x_i}}$

NÄIDE: kaalutud vähimruutude meetod, I

Eesti leibkonnauuring 2012: tarbimiskulud aastas pereliikme kohta ja leibkonna liikmete arv N. Valimi maht 1000.

$$\text{kulud}_i = 4873 - 518N_i + u_i$$

Esineb heteroskedastiivsus.

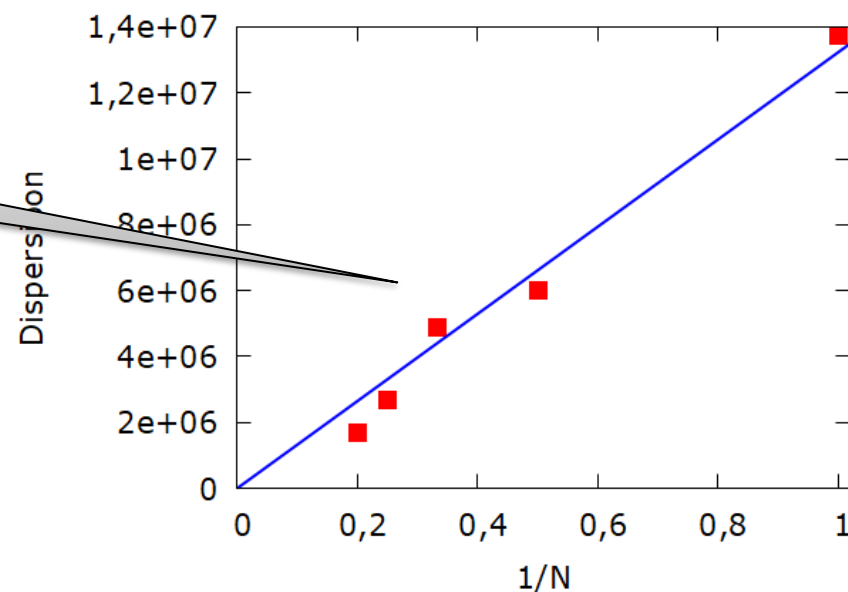
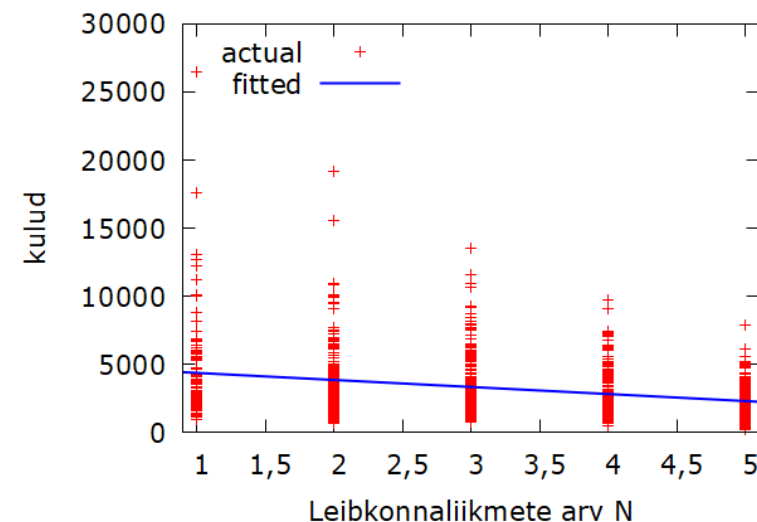
Kuna leibkonna liikmete arv on diskreetne suurus, saame leida tarbimiskulude dispersiooni erineva pereliikmete arvu korral:

N	$\sigma^2 \times 10^{-6}$
1	13,7
2	5,99
3	4,89
4	2,66
5	1,69

Graafik teljestikus $1/N$, σ^2
Võrdeline sõltuvus.

Järelikult $\sigma_i^2 \approx c \frac{1}{N_i}$

Kaal $\frac{1}{\sqrt{\frac{1}{N_i}}} = \sqrt{N_i}$



NÄIDE: kaalutud vähimruutude meetod, II

Leiame WLS mudeli sammhaaval. Kaal $\sqrt{N_i}$

1. Arvutame iga objekti jaoks kaalutud väärtused

$$N_i^* = \sqrt{N_i} \cdot N_i$$

$$\text{kulud}_i^* = \sqrt{N_i} \cdot \text{kulud}_i$$

id	N	kulud	kaal	N*	kulud*
1	2	846	1,414	2,828	1196
2	3	2382	1,732	5,196	4126
3	4	6285	2	8	12571
4	1	3490	1	1	3490
...	...	...	...	...	...

2. Hindame regressioonmudelit kaalutud väärtuste põhjal

$$\text{kulud}_i^* = \beta_1 \sqrt{N_i} + \beta_2 N_i^* + u_i^*$$

WLS tulemus, esitatud
esialgsete tunnuste abil

$$\text{kulud}_i = 4905 - 526N_i + u_i^*$$

Võrdluseks OLS tulemus

$$\text{kulud}_i = 4873 - 518N_i + u_i$$

Praktikas sooritatakse WLS arvutused tarkvara poolt automaatselt, kui kaal ette anda.

WLS ja kaalu sisestamine tarkvara kasutamisel

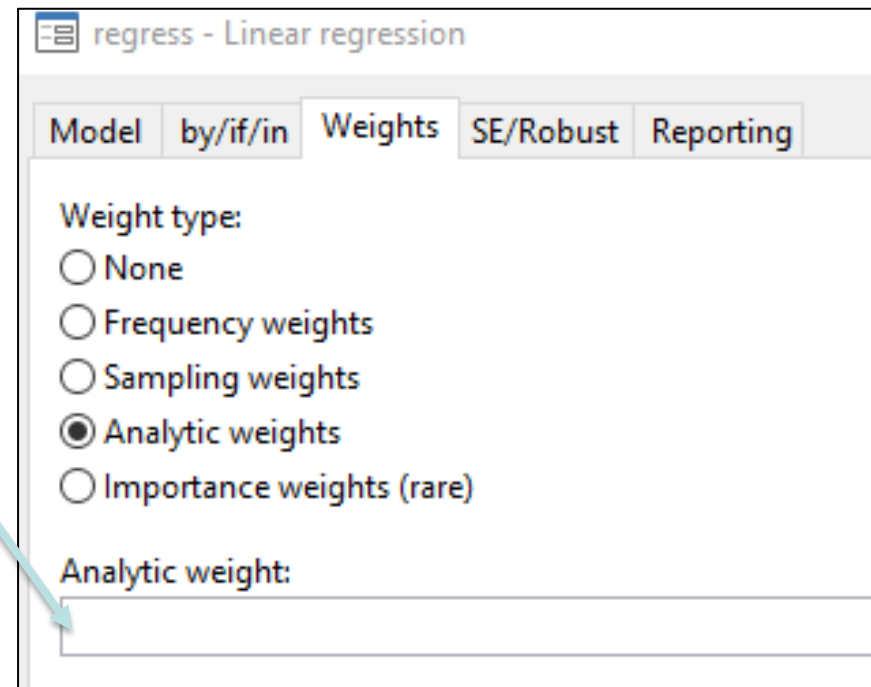
Olgu vealiikmete muutuv dispersioon $Var[u_i | X] = \sigma^2 h_i^2$

Konstantne dispersioon sel juhul $\sigma^2 = \frac{1}{h_i^2} Var[u_i | X]$

Tarkvara kasutamisel tuleb ette anda dispersiooni kaalumiseks vajalik kaal (*weight*).

$$\frac{1}{h_i^2}$$

Näiteks Stata-s lineaarse regressioonmudeli spetsifikatsiooniaknas lehel *Weights*



regress - Linear regression

Model by/if/in Weights SE/Robust Reporting

Weight type:

- ☐ None
- ☐ Frequency weights
- ☐ Sampling weights
- ☒ Analytic weights
- ☐ Importance weights (rare)

Analytic weight:

NÄIDE: pangandussektori analüüs ja WLS

Yeyati ja Micco analüüsisid Ladina-Ameerika pangandussektorit, eraldi kaheksas riigis. Leidsid, et panga risk on konkurentsiga negatiivselt seotud.

Vaatluse all oli palju pankasid erinevatest Ladina-Ameerika riikidest. Näiteks Argentiinast 2337 panka, Brasiiliast 4808 panka, Tšiilist 968 panka jne.

Kasutasid ökonomeetrilist modelleerimist, hindamismeetodiks WLS.

Vaatlused kaaluti läbi vastava panga **varade osakaaluga** (*asset share*).

Kaalumise tulemusena olid suured pangad suurema kaaluga.

1. See tagas, et tulemus kehtib paremini esindusliku panga jaoks.
2. Eeldati, et suuremate pankade korral on mõõtmisvead väiksemad ning nende suurem kaal võimaldab saada täpsemaid tulemusi.

NÄIDE: finantssüsteemi areng ja likviidsus, WLS

Khurana jt analüüsid, kuidas finantssüsteemi areng mõjutab ettevõtete tundlikkust rahavoogude suhtes.

Kasutasid 12 782 ettevõtte andmeid 35 erinevast riigist aastatel 1994-2002.

Modelleeriti sularaha ja turukõlblike väärtpaberite aastase muutuse suhet koguvarasasse.

Seletavateks tunnusteks riiki, aastat ja tööstusharu kirjeldavad fiktiivsed tunnused ning varade logaritm, Tobin Q, rahavood jm.

Ühel juhul oli kaaluks $\frac{1}{\sqrt{N_i}}$, teisel juhul $\frac{1}{\sqrt[3]{N_i}}$, kus N_i oli vaatluste arv riigi kohta.

Teise kaalu kasutamine andis paremaid tulemusi.

Khurana, I., Martin, X., & Pereira, R. (2006). Financial Development and the Cash Flow Sensitivity of Cash. *Journal of Financial and Quantitative Analysis*, 41(4), 787-808.

Khurana, I., Martin, X., & Pereira, R. (2020). Financial Development and the Cash Flow Sensitivity of Cash - CORRIGENDUM. *Journal of Financial and Quantitative Analysis*, 55(3), 1061-1062

WLS kokkuvõte

- Hinnatakse regressioonmudelit, kus esialgsed tunnuste väärtused on asendatud nende kaalutud väärtustega.
- Need vaatlused, mille korral sõltuv tunnus y varieerub rohkem (vealiikmete dispersioon suurem), on väiksema kaaluga ja mõjutavad parameetrite hinnanguid vähem.
- Need vaatlused, mille korral y varieerumine väiksem, on suurema kaaluga, need mõjutavad regressioonmudeli parameetreid rohkem.
- WLS on üldistatud vähimruutude meetodi GLS (*General Least Squares*) erijuht.
 - GLS meetodi korral teisendatakse esialgset mudelit nii, et kehtivad OLS eeldused.
- Kasutada saab ainult siis, kui me teame, kuidas vealiikmete dispersioon muutub.

AGA KUI EI TEA?

Siis rakendatav üldistatud vähimruutude meetod FGLS.

Rakendatav üldistatud vähimruutude meetod FGLS

*Feasible General
Least Squares*

Eeldatakse, et dispersioon sõltub teatud eksogeensetest tunnustest z_j .
Need võivad kokku langeda regressoritega x_j , kuid ei pea.

$$\text{Var}[u_i | Z] = \sigma^2 h_i^2 = \sigma^2 \exp(\alpha_0 + \alpha_1 z_{1i} + \alpha_2 z_{2i} + \dots + \alpha_k z_{ki})$$

Arvu e aste ($\exp$) garanteerib, et dispersioon on positiivne.

Sellist sõltuvust nimetatakse ka **multiplikatiivseks heteroskedastiivsuseks** (*multiplicative heteroscedasticity*).

Siit näeme, et

$$\ln h_i^2 = \alpha_0 + \alpha_1 z_{1i} + \alpha_2 z_{2i} + \dots + \varepsilon_i$$

See mudel kirjeldab dispersiooni muutumist ja kaalude saamiseks hinnatakse seda mudelit.

FGLS etapid

Algne mudel $y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + \dots + u_i \quad (1)$

1. Hinnata mudelit (1) ja leida jääkliikmete hinnangud $\hat{u}_i$
2. Arvutada jääkliikmete hinnangute ruutude logaritmid $\ln h_i^2 = \ln \hat{u}_i^2$
3. Hinnata mudelit (2), kus sõltuvaks tunnuseks on jääkliikmete ruutude logaritmid:

$$\ln h_i^2 = \alpha_0 + \alpha_1 z_{1i} + \alpha_2 z_{2i} + \dots + \alpha_k z_{ki} + \varepsilon_i \quad (2)$$

4. Leida mudelist (2) prognoositud väärtused $\ln h_i^2$
5. Arvutada $h_i^2 = \exp(\ln h_i^2)$ ja seejärel kaalud $1/h_i = 1/\sqrt{h_i^2}$
6. Leida kaalutud väärtused $y_i^* = \frac{y_i}{h_i}, \quad x_{1i}^* = \frac{1}{h_i}, \quad x_{2i}^* = \frac{x_{2i}}{h_i}, \dots$
7. Hinnata OLS abil kaalutud mudelit.

$$y_i^* = \beta_1 x_{1i}^* + \beta_2 x_{2i}^* + \beta_3 x_{3i}^* + \dots + u_i^*$$

FGLS hinnangud

Waldi test

FGLS meetodi korral eeldatakse, et dispersiooni muutumine on kirjeldatud eksogeensete tunnuste z_j komplektiga:

$$\ln h_i^2 = \alpha_0 + \alpha_1 z_{1i} + \alpha_2 z_{2i} + \dots + \alpha_k z_{ki} + \varepsilon_i$$

Kas see eeldus kehtib?

Peale selle mudeli hindamist viiakse läbi Waldi test, kus eemaldatakse mudelist tunnuste komplekt z_j ja testitakse, kas see on lubatud.

$$H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_k = 0$$

Kui tuleb vastu võtta H_0 , siis tunnuste komplekti z_j eemaldamine sellest mudelist on lubatud, mudel ei halvene. Järelikult valitud tunnuste komplekt ei seleta ära dispersiooni muutumist ja FGLS kasutamine selle tunnuste komplektiga pole otstarbekas.

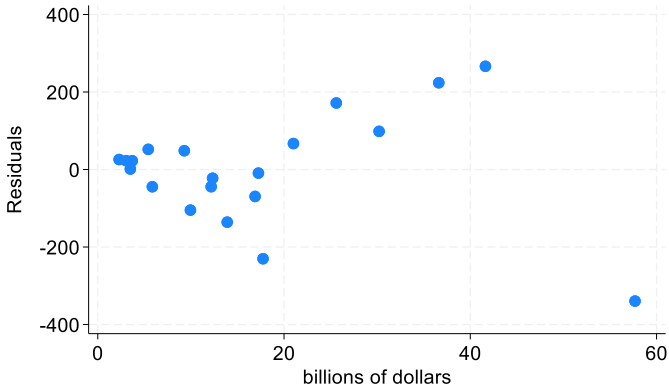
NÄIDE: kasum-käive ja FGLS

20 Saksamaa ettevõtet.

Kasumi sõltuvus käibest

$$profits_i = b_1 + b_2 sales_i + \varepsilon_i$$

Residual versus predictor plot



$$\ln h_i^2 = \alpha_0 + \alpha_1 sales_i + \varepsilon_i$$



Eeldatakse, et dispersioon sõltub regressorist *sales*.

Heteroskedastic linear regression				Number of obs	=	20	
Two-step GLS estimation				Wald chi2(1)	=	3.83	
				Prob > chi2	=	0.0503	
profits	Coefficient	Std. err.	z	P> z	[95% conf. interval]		
profits							
sales	4.922191	2.514533	1.96	0.050	-.0062025	9.850584	
_cons	133.8419	22.65399	5.91	0.000	89.44087	178.2429	
lnsigma2							
sales	.1207585	.0346824	3.48	0.000	.0527822	.1887349	
_cons	7.227046	.7791297	9.28	0.000	5.69998	8.754112	
Wald test of lnsigma2=0: chi2(1) = 12.12					Prob > chi2 = 0.0005		

Waldi test. Võtta vastu H_1 : dispersioon sõltub tunnusest *sales*. Järelikult on FGLS kasutamine otstarbekas.

JÄÄKLIIKMETE AUTOKORRELATSIOON

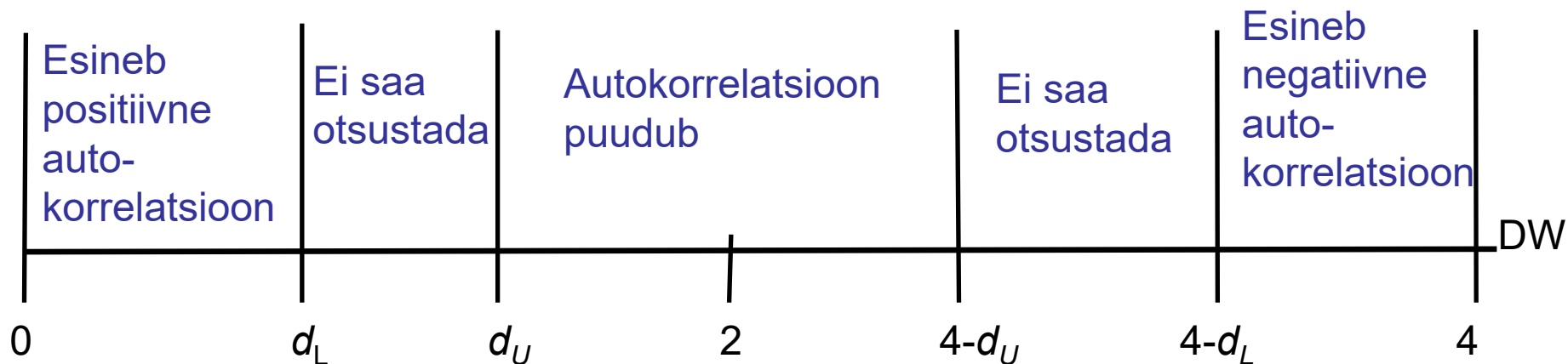
Jääkliikmete autokorrelatsiooni mõju

- Jääkliikmete autokorrelatsioon parameetrite hinnanguid ei mõjuta. Need on ikka **nihketa**.
- Autokorrelatsioon mõjutab parameetrite hinnangute STANDARDVIGU.
 - Parameetrite hinnangud **ei ole enam efektiivsed**.
 - Parameetrite usalduspiirid tulevad valed.
 - Mudeli ja parameetrite olulisuse testimine (F -test ja t -testid) võivad anda valesid tulemusi.
- Järelikult võime teha valesid järeldusi mudelisse kuuluvate või mittekuuluvate tunnuste suhtes.

Jääkliikmete autokorrelatsiooni mõju OLS hinnangutele on samasugune nagu heteroskedastiivsusel.

Jäähäikmete autokorrelatsioonitestimine

- Durbin-Watsoni statistika DW ja selle testimine. 2 kriitilist väärtust D_L ja d_U . DW annab teada ainult 1. järku autokorrelatsioonist



- Breusch-Godfrey LM test. Saab testida autokorrelatsiooni olemasolu järguni r

$$\hat{u}_t = \gamma_1 + \gamma_2 x_{2t} + \gamma_3 x_{3t} + \rho_1 \hat{u}_{t-1} + \rho_2 \hat{u}_{t-2} + \dots + \rho_r \hat{u}_{t-r} + v_t$$

$H_0: \rho_1 = \rho_2 = \dots = \rho_r = 0$ autokorrelatsioon puudub

- Jäähäikmete korrelogramm ja Q-statistik.

Mida teha, kui esineb autokorrelatsioon

- Katsetada erineva kujuga mudeleid.
- Põhjuseks võib olla mõne olulise tunnuse puudumine mudelist: lisada mudelisse tunnuseid.
- 1. järku autokorrelatsiooni saab eemaldada spetsiaalse protseduuriga:
 - Cochrane-Orcutt protseduur
 - Prais-Winsteni protseduur
 - Hildreth-Lu protseduur
- Kohandatud Newey-West standardvigade kasutamine.
 - Kasutatavad suurte valimite korral.
- Kasutada dünaamilist regressioonmudelit.

Jääkide autokorrelatsiooni võimalik põhjus

Sargan jt: jääkliikmete autokorrelatsioon tuleneb sellest, et mudeli dünaamika ei ole korrektselt spetsifitseeritud.

Näiteks lineaarne mudel $y_t = b_1 + b_2x_{2t} + b_3x_{3t} + \dots + b_kx_{kt} + u_t$

- Eeldab seda, et esineb vaid regressorite **kohene** mõju sõltuvale tunnusele.
- Jääkliikmete autokorrelatsioon võib tulla sellest, et aegrea y dünaamikat ei kirjelda see mudel korrektselt.
- Sisuliselt on tegemist staatilise mudeliga: kõikide tunnuste väärtused on ühel ja samal ajaperioodil (-hetkel).

Probleemi lahendab dünaamilise mudeli kasutamine.

Dünaamilised mudelid

Staatiline mudel modelleerib vaid kohest mõju $y_t = b_1 + b_2x_{2t} + b_3x_{3t} + u_t$

Dünaamiline mudel: y väärtus ajahetkel t võib sõltuda ka y enda või regressorite väärtustest eelmistel ajaperioodidel.

Regressorite viitajad:

jaotatud viitajaga

(distributed lag) mudel

$$y_t = \alpha + \beta_0x_{2t} + \beta_1x_{2t-1} + \beta_2x_{2t-2} + \gamma_0x_{3t} + \gamma_1x_{3t-1} + \gamma_2x_{3t-2} + u_t \quad \text{DL}$$

Sõltuva tunnuse viitajad:

autoregressiivne mudel

$$y_t = \alpha + \varphi_1y_{t-1} + \beta x_{2t} + \gamma x_{3t} + u_t \quad \text{AR(1)}$$

Sõltuva tunnuse ja regressorite viitajad:

jaotatud viitajaga autoregressiivne ehk

ARDL mudel (autoregressive distributed lag)

$$y_t = \alpha + \varphi_1y_{t-1} + \beta_0x_t + \beta_1x_{t-1} + u_t \quad \text{ARDL(1,1)}$$

Võib lisada ka kõrgemat järku viitaegu.

Dünaamilisi mudeleid kasutatakse mitte ainult jääkide autokorrelatsiooni eemaldamiseks, vaid ka majandusloogikast lähtudes pikaajalise mõju analüüsimiseks.

NÄIDE: inflatsioon ja raha pakkumine, jaotatud viitajaga mudel

Raha pakkumise mudel USA-s 1955 1 kv kuni 1969 IV kv

P tarbijahindade muutus aastas, %

M raha pakkumise muutus aastas, %

$$P_t = -0,146 + \sum_{i=0}^{20} m_i M_{t-i}$$

Pikaajaline mõju $\sum_{i=0}^{20} m_i = 1,031$

Raha pakkumise suurenemine mingis kvartalis 1% võrra suurendas pikaajaliselt (20 kvartali ehk 5 aasta jooksul) tarbijahindasid ca 1%.

Koefitsient	
m0	0,041
m1	0,034
m2	0,030
m3	0,029
m4	0,030
m5	0,033
m6	0,037
m7	0,042
m8	0,048
m9	0,054
m10	0,059

Koefitsient	
m11	0,065
m12	0,069
m13	0,072
m14	0,073
m15	0,072
m16	0,069
m17	0,062
m18	0,053
m19	0,039
m20	0,022

Mõjukordajad

Jaotatud viitajaga, DL $y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \dots + \beta_k x_{t-k} + u_t$

β_0 **lühiajaline mõjukordaja** (*short-run multiplier*). Kui x kasvab perioodil t ühe ühiku võrra, siis y kasvab samal perioodil β_0 ühikut.

Kogumõju peale k perioodi näitab **pikaajaline mõjukordaja** (*long-run multiplier*)

$\beta_{LR} = \sum_{i=0}^k \beta_i$ Kui x kasvab perioodil t ühe ühiku võrra, siis y kasv kõigi järgnevate perioodide vältel kokku on β_{LR} ühikut.

Autoregressiivne jaotatud viitajaga $y_t = \alpha + \varphi_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + u_t$ ARDL(1,1)

Pikaajaline mõjukordaja $\beta_{LR} = \frac{\beta_0 + \beta_1}{1 - \varphi_1}$  Arvestab seda, et y_{t-1} on ka mõjutatud x_{t-1} poolt.

ARDL(p,k) mudeli korral $\beta_{LR} = \frac{\sum_{i=0}^k \beta_i}{1 - \sum_{j=1}^p \varphi_j}$

INSTRUMENTTUNNUSED

Eksogeensed ja endogeensed regressorid

$$y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + \dots + \varepsilon_i$$

Endogeense tunnuse väärtused on määratud mudeli siseselt.

Eksogeense tunnuse väärtused on määratud mudelist väljaspool.

Kui regressoriks olev tunnus (mudelis paremal pool) on korrelatsioonis juhusliku liikmega, siis see regressor on endogeenne.

Endogeense regressori korral	$Cov(x_i, \varepsilon_i) \neq 0$
------------------------------	----------------------------------

Eksogeense regressori korral	$Cov(x_i, \varepsilon_i) = 0$
------------------------------	-------------------------------

Mis võib tekitada regressori endogeensuse?

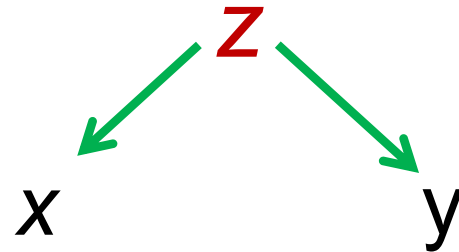
- Varjatud ehk latentse tunnuse olemasolu.
- Oluline tunnus on mudelist välja jäänud.
- Mõõtmisvead.
- Simultaansus ja vastastik-kausaalsus

Latentne tunnus

Meie modelleerime

$$x \longrightarrow y$$

Aga tegelikult eksisteerib latentne tunnus z , mida me pole mõõtnud ja mis mõjub mõlemale:



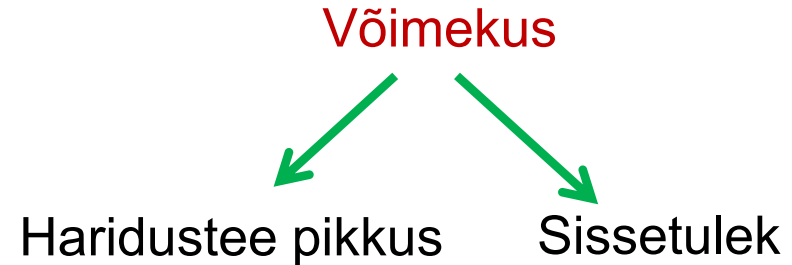
Latentne tunnus, näited

Modelleeritakse

Haridustee pikkus → Sissetulek

Ei arvestata võimekust (latentne tunnus).

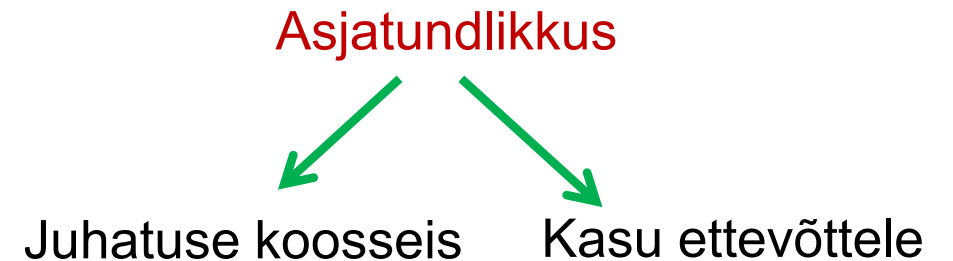
Tegelik seos



Analoogne probleem ettevõtte juhatuse analüüsimisel.

Juhatuse koosseis → Kasu ettevõttele

Ettevõtted, kes sellest asjatundlikkusest rohkem kasu saavad, otsivad tõenäolisemalt juhatuse liikmeid, kel on vajalik spetsiifiline taust.



Siis OLS ülehindab millegi mõju sõltuvale tunnusele.

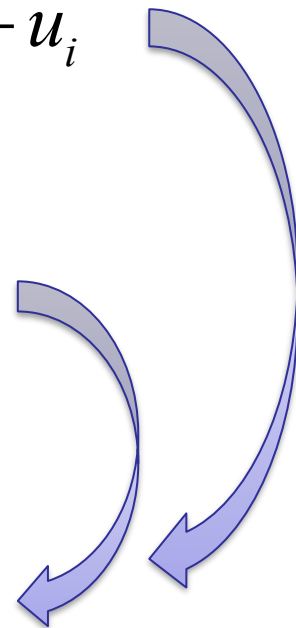
Välja on jäetud oluline tunnus

Tegelik seos $y_i = \beta_1 + \beta_2 x_i + \beta_3 z_i + u_i$

Meie hindame $y_i = \gamma_1 + \gamma_2 x_i + \varepsilon_i$ $\varepsilon_i = \beta_3 z_i + u_i$

Kui x ja z on omavahel mingil määral seotud $Cov(x_i, z_i) \neq 0$

Tulemus $Cov(x_t, \varepsilon_t) \neq 0$



Mõõtmisvead

Näiteks tegelik seos on $y_i = \beta_1 + \beta_2 x_i^* + \varepsilon_i$

Regressori tegelike väärtuste x_i^* asemel on meil mõõdetud väärtused x_i , mis on mõõtmisveaga:

$$x_i = x_i^* + v_i$$

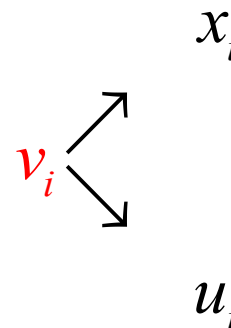
↙ Mõõtmisviga

Seega me hindame mudelit $y_i = \beta_1 + \beta_2 x_i + u_i$

Tegelik seos mõõdetava suuruse x kaudu

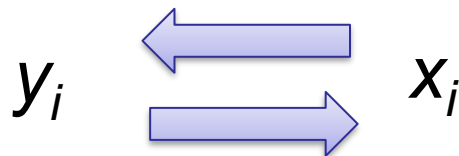
$$y_i = \beta_1 + \beta_2 x_i^* + \varepsilon_i = \beta_1 + \beta_2 (x_i - v_i) + \varepsilon_i = \beta_1 + \beta_2 x_i - \beta_2 v_i + \varepsilon_i = \beta_1 + \beta_2 x_i + u_i$$

$$u_i = \varepsilon_i - \beta_2 v_i$$


$$\Rightarrow \text{Cov}(x_i, u_i) \neq 0$$

Simultaansus ja vastastik-kausaalsus

Vastastik-kausaalsusest tingitud endogeensuse probleem võib ilmneda seetõttu, et mitte üksnes x_i ei avalda mõju y_i -le, vaid samal ajal ka y_i avaldab mõju regressorile x_i



Näiteks tarbimine elaniku kohta y ja sissetulek elaniku kohta x :



$$y_t = \beta_1 + \beta_2 x_t + \varepsilon_t$$

Aga sissetulek pole eksogeenne, vaid on määratud samasusega

Sissetulekutase (x) $\equiv$ kogutarbimine (y) + investeeringud elaniku kohta (z)

$$x_t = \gamma_1 + \gamma_2 y_t + \gamma_3 z_t + u_t$$



Regressori endogeensuse tagajärjed

Kui eksogeensuse eeldus pole täidetud, siis OLS parameetrite hinnangud

- võivad olla nihkega;
- ei ole efektiivsed;
- ei ole mõjusad.

Standardvigade hinnangud:

- võivad olla nihkega;
- ei ole mõjusad.

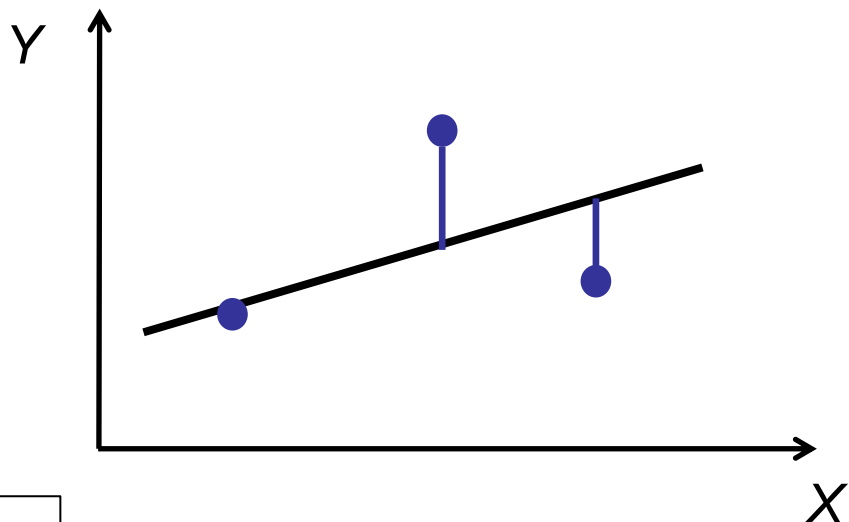
Kui eksogeensuse eeldus pole täidetud

Kui näiteks tunnuse x_2 ja juhusliku liikme ε vahel on seos, siis me ei saa tunnuse x_2 mõju sõltuvalle tunnusele y puhtalt eraldada.

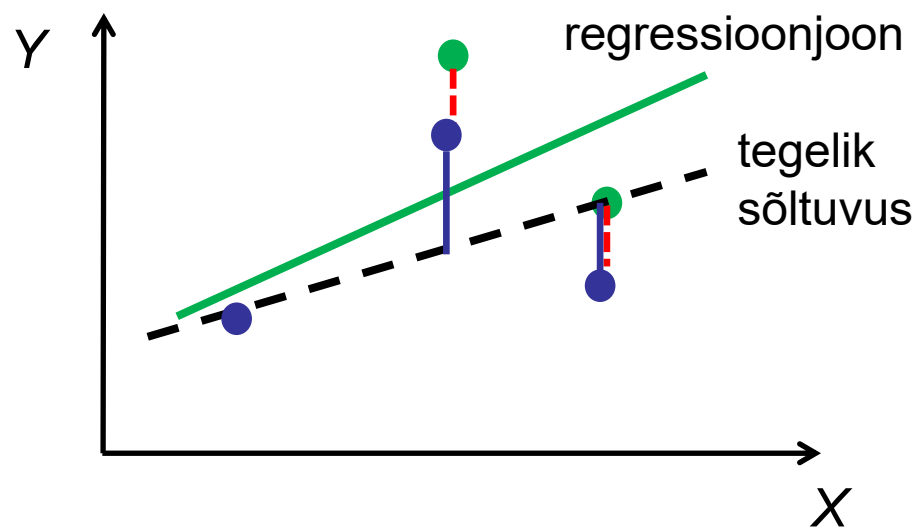
$$y_i = \hat{\beta}_1 + \hat{\beta}_2 x_{2i} + \varepsilon_i$$

Kui näiteks x_2 ja ε vahel on positiivne korrelatsioon, siis me saame parameetri β_2 hinnangu suurema.

$$\text{Cov}(\varepsilon_i, X_{2j}) = 0$$



$$\text{Cov}(\varepsilon_i, X_{2j}) > 0$$



Demo

Parameetrite hinnangud on nihkega

Mida teha, kui eksogeensuse eeldus pole täidetud?

Instrument variables, IV

Kasutada instrumenttunnuseid.

Instrumenttunnuste kasutamist vaadeldi esmakordselt raamatu Wright, Philip G. 1928. The Tariff on Animal and Vegetable Oils. New York: Macmillan lisas Appendix B.

Rasvainete nõudlus- ja pakkumisfunktsiooni polnud võimalik kindlaks teha ja tuli kasutada instrumenttunnuseid.

Instrumentide kasutamine finantsökonomeetrias

Instrumenttunnuseid kasutatakse eriti uuringutes, mis käsitlevad

- kapitalistruktuuri
 - Finantsvõimendus (*leverage*) võib olla endogeenne, sest ettevõtte väärtus mõjutab ka finantseerimisotsuseid.
- likviidsust ja aktsiate tootlust
 - Likviidsus ja tootlus võivad olla samaaegselt määratud.
- investorite käitumist
 - Investorite turumeeleolu (sentiment) võib olla endogeenne.
- fondide rahavooge
 - Fondi rahavood ja tootlus mõjutavad teineteist.

Näiteid instrumentide kasutamisest

Erinevad artiklid ajakirjadest
Journal of Finance,
Journal of Financial Economics,
Review of Financial Studies,
Journal of Financial,
Quantitative Analysis
perioodil 1997-2012.

Wang, C. J. (2015). Instrumental
variables approach to correct for
endogeneity in finance. Handbook of
financial econometrics and statistics,
2577-2600.

Table 95.1 Finance research that uses instrumental variable methods

Capital structure/leverage ratio	Debt covenants
Faulkender and Petersen (RFS 2006)	Dennis et al. (JFQA 2000)
Yan (JFQA 2006)	Chen et al. (JF 2007)
Molina (JF 2005)	Macro/product market
Johnson (RFS 2003)	Thorsten et al. (JFE 2000)
Desai et al. (JF 2004)	Campello (JFE 2006)
Harvey et al. (JFE 2004)	Garmaise (RFS 2008)
Agency/ownership structure/governance	Financial institutions
Bitler et al. (JF 2005)	Ljungqvist et al. (JF 2006)
Ortiz-Molina (JFQA 2006)	Berger et al. (JFE 2005)
Palia (RFS 2001)	Microstructure
Cho (JFE 1998)	Brown et al. (JF 2008)
Wei et al. (JFQA 2005)	Conrad et al. (JFE 2003)
Daines (JFE 2001)	Kavajecz and Odders-White (RFS 2001)
Coles et al. (JFE 2012)	Diversification/acquisition
Wintoki et al. (JFE forthcoming)	Campa and Kedia (JF 2002)
Pricing of public offering	Hsieh and Walkling (JFE 2005)
Cliff and Denis (JF 2004)	Venture capital/private equity
Lee and Wahal (JFE 2004)	Gompers and Lerner (JFE 2000)
Lowry and Shu (JFE 2002)	

Instrumenttunnuste kasutamine põhjuslike seoste analüüsimisel

Perioodil 2003-2014 ajakirjades
Journal of Finance,
Journal of Financial Economics,
Review of Financial Studies
ilmunud artiklid, kus põhjusliku seose hindamisel kasutati
instrumenttunnuseid IV (255 artiklit).

Klassifitseeris artiklid kolmeks sõltuvalt sellest,
kuidas OLS ja IV hinnangud erinesid.

67,1% OLS ülehindas mõju
Nn jaatav endogeensus (*affirmative endogeneity*)

18,0% OLS alahindas mõju
Nn korrigeeriv endogeensus (*corrective endogeneity*)

14,9% OLS ja IV hinnangute erinevust ei saanud määrata
(*unclear*)

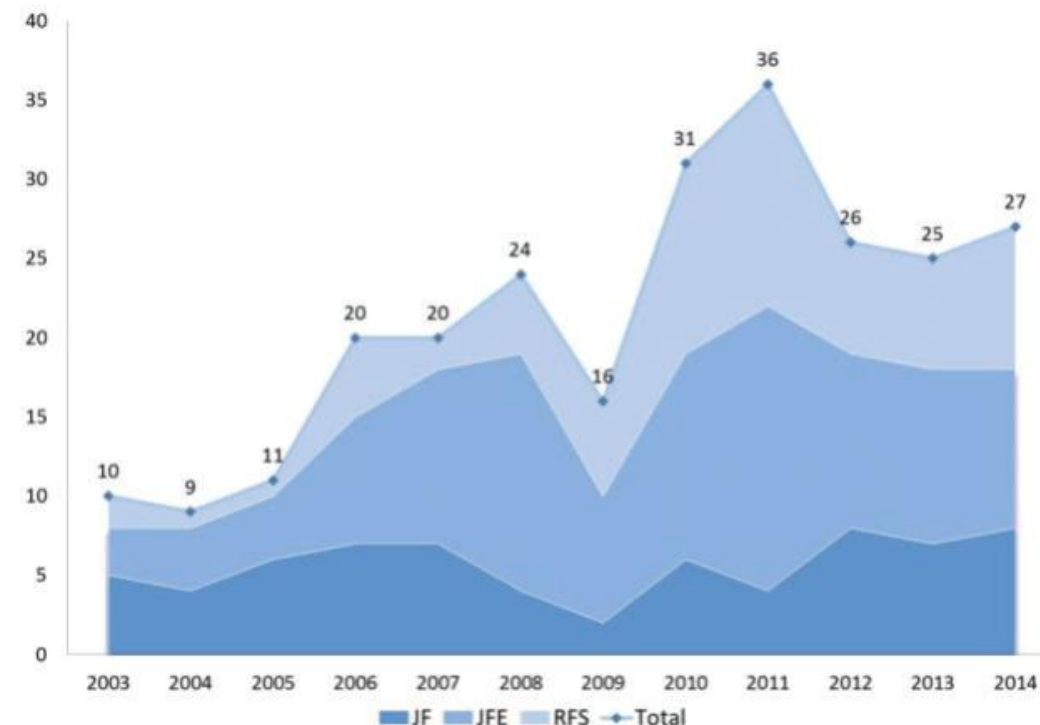


Figure 1.
Number of published papers.

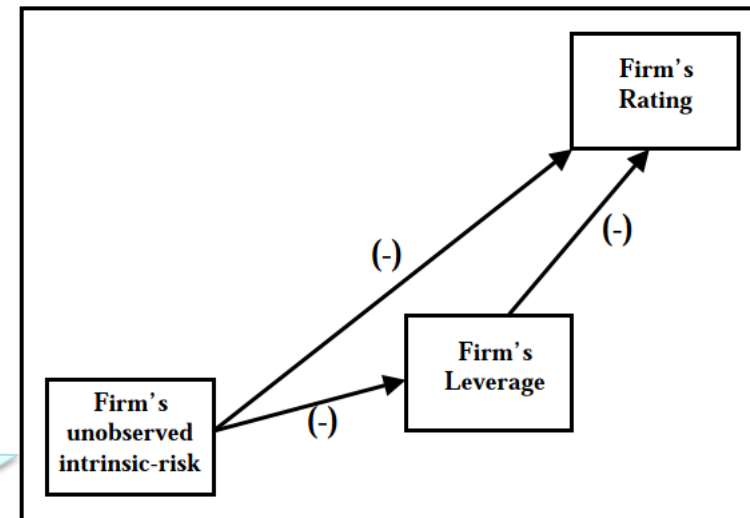
NÄIDE: finantsvõimenduse efekt ettevõtte reitingule ja instrumendid

Molina analüüsis, kuidas finantsvõimendus (*leverage*) mõjutab ettevõtte reitingut. 2678 vaatlust (ettevõtte-aasta).

Hinnatav mudel $Rat_i = \alpha + \beta_1 Lev_i + \beta_2 \mathbf{X}_i + \varepsilon_i$

kus $\mathbf{X}_i$ on kontrolltunnuste vektor (ettevõtte suurus, kasumlikkus jt).

Finantsvõimendus Lev on endogeenne, sest mittevaadeldavad tunnused mõjutavad nii finantsvõimendust kui ka reitingut.



Instrumentideks näitajad, mis mõjutavad finantsvõimendust, kuid ei mõjuta ettevõtte reitingut:

- ettevõtte väärtuse varasemate aastate hinnangud (*history of past market valuations*)
- ettevõtte maksu piirmäär (*firm's marginal tax rate*).

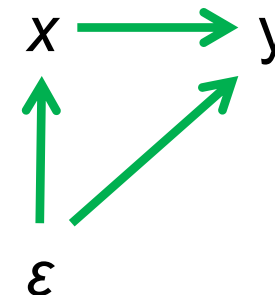
Tulemus: finantsvõimenduse efekt ettevõtte reitingule on instrumentide kasutamisel ca 3 korda suurem kui endogeensuse ignoreerimisel. OLS alahindas finantsvõimenduse mõju.

Molina, C. A. (2005). Are firms underleveraged? An examination of the effect of leverage on default probabilities. *The Journal of Finance*, 60(3), 1427-1459.

Mis on instrument?

$$y_i = \beta_1 + \beta_2 x_i + \varepsilon_i$$

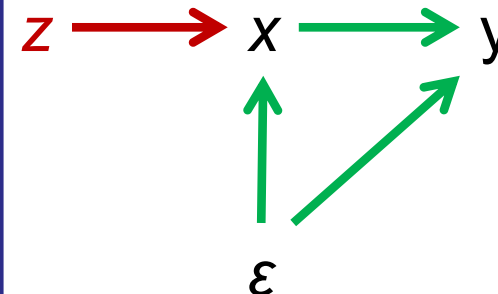
Endogeenne regressor $Cov(x_i, \varepsilon_i) \neq 0$



Eesmärgiks on leida parameetrite β_1 and β_2 jaoks mõjusad hinnangud.

Instrument: uus tunnus **z**, mis peab rahuldama tingimusi

- | | | |
|-----|-------------------------------|---|
| (1) | $Cov(z_i, \varepsilon_i) = 0$ | eksogeensus |
| (2) | $Cov(z_i, x_i) \neq 0$ | kohasus, st on seotud endogeense regressoriga |



Kui need kaks tingimust on täidetud, siis on **z** tunnuse **x** instrumenttunnus.
Meetod IV (*instrumental variables method*).

Kas tingimused (1) ja (2) on rahuldatud?

Instrumendi **eksogeensuse tingimust** (*exogeneity condition*) (1)

$$\text{Cov}(z_i, \varepsilon_i) = 0$$

statistiliselt kontrollida ei saa, sest juhuslikud liikmed ε pole vaadeldavad.
Kontrollimiseks loogiline arutluskäik.

Instrumendi **kohasuse tingimust** (*relevance condition*) (2)

$$\text{Cov}(z_i, x_i) \neq 0$$

st kas instrument on seotud endogeense regressoriga, saab kontrollida,
hinnates regressioonmudelit

endogeenne regressor

$$x_i = \delta_1 + \delta_2 z_i + w_i$$

$H_0 : \delta_2 = 0, \quad H_1 : \delta_2 \neq 0$

instrument

Kui H_0 ümber lükatud, siis on seos olemas.

Mitmene regressioonimudel

$$y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + \varepsilon_i$$

$$\text{Cov}(x_{2i}, \varepsilon_i) \neq 0 \quad x_2 \text{ endogeenne}$$

$$\text{Cov}(x_{3i}, \varepsilon_i) = 0 \quad x_3 \text{ eksogeenne}$$

Kaks endogeenset tunnust: x_2 ja y .

Üks eksogeenne tunnus: x_3 .

Oluline on eristada **endogeenseid** ja **eksogeenseid** tunnuseid.

Seepärast edaspidi uued tähistused: **endogeensed** $y_1, y_2, \dots$
eksogeensed $z_1, z_2, \dots$

$$y_{1i} = \beta_1 + \beta_2 y_{2i} + \beta_3 z_{1i} + \varepsilon_i$$

Mitmene regressioonimudel ja üks instrument

$$y_{1i} = \beta_1 + \beta_2 y_{2i} + \beta_3 z_{1i} + \varepsilon_i \quad (1)$$

Kas eksogeenne tunnus z_1 võib olla endogeense y_2 instrument?

Ei või, sest instrument ei tohi olla mudelis (1).

Vaja veel üht eksogeenset tunnust z_2 , mis ei esine mudelis (1) ja rahuldab instrumendi tingimusi:

$$\text{Cov}(z_{2i}, \varepsilon_i) = 0$$

$$\text{Cov}(z_{2i}, y_{2i}) \neq 0$$

Redutseeritud ja struktuurne kuju

$$y_{1i} = \beta_1 + \beta_2 y_{2i} + \beta_3 z_{1i} + \varepsilon_i \quad (1)$$

Kasutame instrumenti z_2 . Kas see on kohane? $Cov(z_{2i}, y_{2i}) \neq 0$?

Kontrollimiseks hindame mudelit $y_{2i} = \delta_1 + \delta_2 z_{1i} + \delta_3 z_{2i} + w_i \quad (2)$

Kui δ_3 on statistiliselt oluline, siis z_2 on y_2 jaoks kehtiv instrument.

Mudeliga (2) on endogeenne tunnus y_2 avaldatud eksogeensete tunnuste kaudu. See on võrrandi **redutseeritud kuju**.

Esialgne mudel (1), kus regressoriteks nii endogeenne kui eksogeenne: **struktuurne kuju**. See väljendab uuritava nähtuse struktuuri.

Ühel endogeensel tunnusel mitu instrumenti

$$y_{1i} = \beta_1 + \beta_2 y_{2i} + \beta_3 z_{1i} + \varepsilon_i \quad (1)$$

Kas ühe endogeense regressori jaoks võib olla mitu instrumenti?

Võib küll.

1. Kumbki ei tohi esineda mudelis (1).
2. Peavad olema täidetud eksogeensuse tingimused: $Cov(z_{2i}, \varepsilon_i) = 0, Cov(z_{3i}, \varepsilon_i) = 0$
3. Peavad olema korrelatsioonis vastava endogeense regressoriga y_2

Tingimuse 3 kontrollimiseks mudel redutseeritud kujul

$$y_{2i} = \delta_1 + \delta_2 z_{1i} + \delta_3 z_{2i} + \delta_4 z_{3i} + w_i \quad (2)$$

Kitsenduste F -test. Kui $H_0: \delta_3 = \delta_4 = 0$ on ümber lükatud, siis kõlbavad.

Identifitseerimine

Üldiselt võib endogeenseid regressoreid olla mitu.

Endogeensete regressorite arv k : $y_1 \dots y_k$

Parameetrite arv k : $\beta_1 \dots \beta_k$

Instrumentide arv m : $z_1 \dots z_m$

Täpselt identifitseeritud $m = k$ Instrumentide ja endogeensete arv ühesugune

Üleidentifitseeritud $m > k$ Instrumente rohkem

Alaidentifitseeritud $m < k$ Instrumente vähem kui endogeenseid

Parameetreid ei saa leida, kui on alaidentifitseeritud.

Hindamismeetodid instrumenttunnuste kasutamisel

- Kahe-etapiline vähimruutude meetod 2SLS (*Two stage least squares*)

Simultaansete võrrandite korral

- Piiratud informatsiooniga suurima tõepära meetod LIML (*Limited-information maximum likelihood*)
- Üldistatud momentide meetod GMM (*Generalized method of moments*)

Kahe-etapiline vähimruutude meetod 2SLS

Struktuurne kuju $y_{1i} = \beta_1 + \beta_2 y_{2i} + \beta_3 z_{1i} + \varepsilon_i \quad (1)$

1. Avaldada endogeenne regressor y_2 kõigi eksogeensete tunnuste kaudu (redutseeritud kuju) ja leida selle hinnangud. $\hat{y}_{2i} = \hat{\delta}_1 + \hat{\delta}_2 z_{1i} + \hat{\delta}_3 z_{2i} + \hat{\delta}_4 z_{3i}$
2. Asendada leitud hinnangud struktuursesse võrrandisse (1) ja leida parameetrite hinnangud. $y_{1i} = \beta_1 + \beta_2 \hat{y}_{2i} + \beta_3 z_{1i} + u_i$

OLS meetodit rakendatakse 2 korda: *two stage least squares* (2SLS).

NÄIDE: CCAPM mudel

Tarbimisega laiendatud finantsvarade hindamise mudel CCAPM
(*consumption-based capital asset price model*).

CAPM mudelis riskantne väärtpaber suurendab määramatust investori rikkuses.
CCAPM mudeli korral suurendab riskantne väärtpaber määramatust investori **tarbimises**. See mudel näitab, kuidas aktsiaturg on seotud tarbimise kasvuga.

$$y_t = \alpha + \gamma cg_t + \varepsilon_t$$

y_t tulumäär (logaritmiline)

cg_t agregeeritud tarbimise kasv, logaritmitud

γ on suhtelise riskikartlikkuse koefitsient (*coefficient of relative risk aversion*).

Oodatava tulumäära standardhälve on γ korda suurem kui oodatava tarbimise kasvu standardhälve.

Campbell, J. Y., Lo, A. W., and MacKinlay, A. C. (1997). The Econometrics of Financial Markets. Princeton University Press.

Geyer, A. (2021) Basic Financial Econometrics.

Lühidalt CCAPM mudeli kohta loe nt Investopedia <https://www.investopedia.com/terms/c/ccapm.asp>

NÄIDE: CCAPM mudel, OLS

Andmed S&P500 kohta aastatel 1939-2004 (Geyer).

Oodatav logaritmiline tulumäär $y_t = E_{t-1}[\ln(1 + R_t)] + u_t$ R_t tulumäär

Oodatav tarbimise kasv $cg_t = E_{t-1}[\Delta \ln C_t] + v_t$ C_t agregeeritud tarbimine

Mõlema tunnuse määramisel esinevad mõõtmisvead.

CCAPM mudel $y_t = \alpha + \gamma cg_t + \varepsilon_t$

OLS hindamise tulemus $y_t = -0,019 + 3,64 cg_t + \varepsilon_t$

OLS-i aga kasutada ei tohiks, sest CCAPM vealiige ε_t on korrelatsioonis tarbimise kasvuga, on sellega seotud läbi mõõtmisvea v_t

$$\varepsilon_t = u_t - \gamma v_t \quad Cov(cg_t, \varepsilon_t) \neq 0$$

Endogeensus on tingitud mõõtmisvigadest.

Instrumentideks võib kasutada mingit tarbimisega korrelatsioonis oleva suuruse viitaega.

Geyer, Campbell pakkusid: eelmise perioodi intressimäär i_{t-1}
dividendide ja hinna suhe dp_{t-1}

NÄIDE: CCAPM mudel, IV meetod

$$y_t = \alpha + \gamma cg_t + \varepsilon_t$$

Instrumenttunnuste
meetodil

Kuvatakse struktuurse kuju
aruanne (2SLS teine etapp)

y	Coefficient	Std. err.	z	P> z
cg	7.926	2.662	2.98	0.003
_cons	-0.121	0.066	-1.82	0.069

Instrumented: cg
Instruments: L.i L.dp

OLS-i korral oli 3,64

Endogeenne tunnus

Lisaks võib lasta kuvada ka
redutseeritud kuju aruande
(2SLS esimene etapp)

cg	Coefficient	Std. err.	t	P> t
i				
L1.	.2036788	.0668061	3.05	0.003
dp				
L1.	.0033859	.0061172	0.55	0.582
_cons	.032856	.0202778	1.62	0.110

Instrumentid

Nõrk instrument

Endogeensuse testimise vajalikkus

Kui efektiivne hinnang ei ole mõjus, on parem kasutada mõjusat ebaefektiivset hinnangut.

OLS vs IV meetod:

OLS hinnangud on **efektiivsemad** kui IV hinnangud: $\text{Var}(b_{OLS}) < \text{Var}(b_{IV})$

Aga OLS hinnangud ei ole alati mõjusad.

IV hinnangud on **mõjusad**.

Kui kõik regressorid on eksogeensed, siis

- OLS meetodil leitud hinnangud on mõjusad.

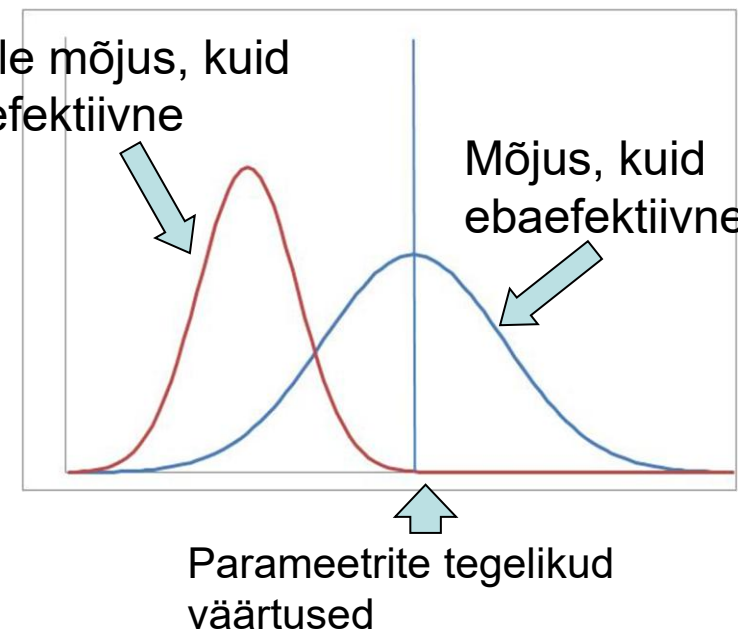
Kui mõned regressorid ei ole eksogeensed (mõni on endogeenne), siis

- OLS meetodil leitud hinnangud ei ole mõjusad.

Eksogeensete regressorite korral tuleb eelistada OLS meetodit.
Kui mõned endogeensed, siis IV meetod.

Ei ole mõjus, kuid on efektiivne

Mõjus, kuid ebaefektiivne

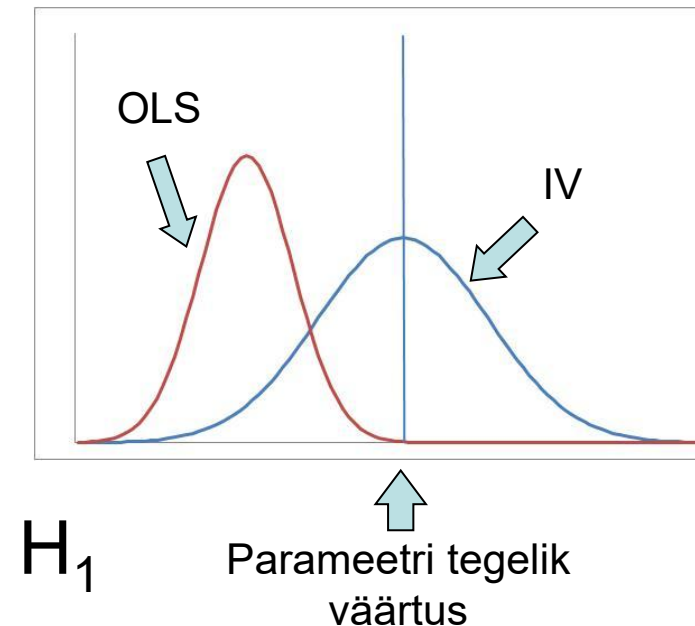
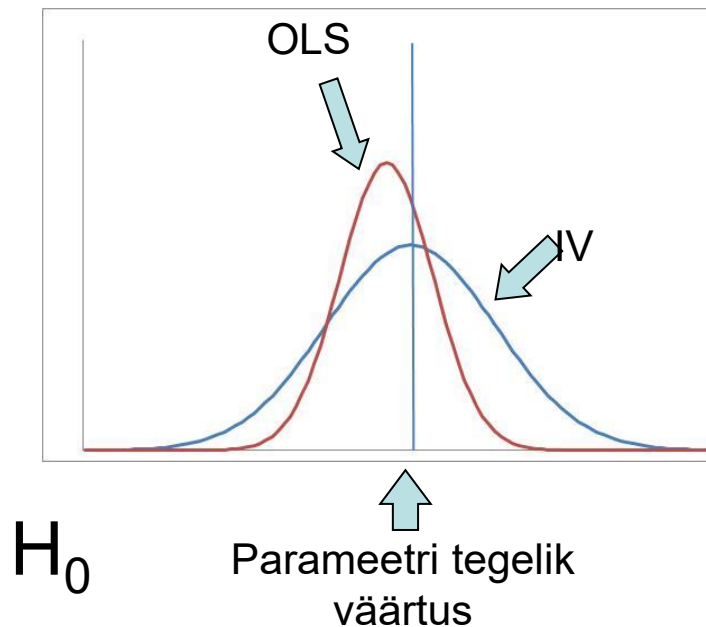


Endogeensust on vaja testida: kas IV meetodi kasutamine on põhjendatud.
Kasutatakse Durbin-Wu-Hausmani spetsifikatsiooni testi.

Durbin-Wu-Hausmani spetsifikatsiooni test

Hausmani testi idee seisneb kahe erineva hinnangfunktsiooni abil saadud hinnangute võrdlemises, kus üks on efektiivne ja teine ei ole.

1. Mudeli hindamiseks kasutatakse OLS-i ja instrumenttunnuste meetodit.
2. Testitakse, kas need hinnangfunktsioonid annavad statistiliselt oluliselt erinevad hinnangud.
3. Kui hinnangud **oluliselt ei erine** (H_0), siis nii IV kui ka OLS hinnangud on mõjusad. Siis tuleb kasutada OLS-i, sest see on efektiivne.
4. Kui H_0 on ümber lükatud, siis OLS hinnangud ei ole mõjusad ja esineb endogeensus. Siis kasutada instrumenttunnuste meetodit IV.



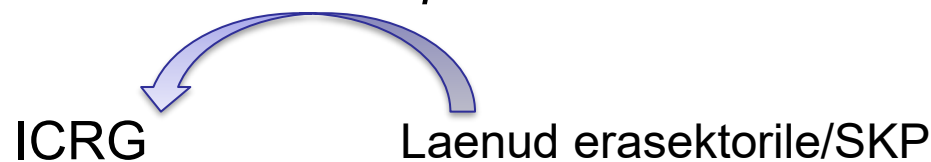
NÄIDE: kas finantssüsteemi areng vähendab korruptsiooni

Sõltuv tunnus korruptsiooni indeks ICRG

International Country Risk Guide indicator of corruption

Vaatluse all 410 riiki.

Endogeensuse probleem:
olulised tunnused võivad
puududa.



Instrumendiks võeti riigi **õigussüsteem**



finantssüsteem

4 võimalikku

- Inglise tsiviilõigus
- Prantsusmaa tsiviilõigus
- Saksa tsiviilõigus
- Skandinaavia tsiviilõigus

Tulemused: OLS alahindas

	OLS	IV
Constant	−2.7885*** (0.1608)	−2.5948*** (0.1761)
Bank lending to the private sector	−0.0051*** (0.0008)	−0.0094*** (0.0007)
Durbin–Wu–Hausman test statistic (p-value)		8.688 (0.0032)

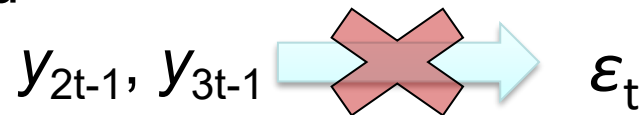
Viitajad instrumentidena

Aegridade kasutamisel võetakse instrumentideks tihti regressorite viitajad.

$$y_{1t} = \beta_1 + \beta_2 y_{2t} + \beta_3 y_{3t} + \varepsilon_t$$

Kui y_2 , y_3 on endogeensed regressorid, siis

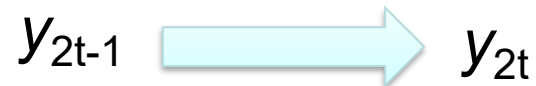
instrumendid



$$Cov(y_{2t-1}, \varepsilon_t) = 0$$

$$Cov(y_{3t-1}, \varepsilon_t) = 0$$

Eksogeensuse
tingimus on täidetud



$$Cov(y_{2t-1}, y_{2t}) \neq 0$$



$$Cov(y_{3t-1}, y_{3t}) \neq 0$$

Kohasuse tingimus on
täidetud, kui esineb
autokorrelatsioon

Endogeensuse probleemi lahendamiseks kasutatakse dünaamilist mudelit.

Miks kasutatakse instrumente harva?

- Sobivate instrumentide leidmine on raske.
- Instrumendid võivad olla nõrgad (*weak or poor instruments*).
 - On küll eksogeensed, kuid nõrgalt seotud endogeense regressoriga.
- Teiselt poolt, paljud eksogeensed tunnused, mis on seotud endogeensete regressoritega, mõjutavad samal ajal otseselt ka sõltuvat tunnust. Aga instrument seda teha ei tohi.
- Väikeste valimite korral instrumentide kasutamine hästi ei sobi.
 - Lõpliku valimi korral ei ole instrumentide omadused korrektsed.
 - Asümptootilised omadused (suure valimi korral) on sobivad.