Norges teknisk-naturvitenskapelige universitet
Institutt for samfunnsøkonomi

EKSAMENSOPPGAVE I SØK3001

ØKONOMETRI I

ECONOMETRICS I

Faglig kontakt under eksamen: Kåre Johansen
Tlf.: 9 19 36

Eksamensdato: Torsdag 7. juni 2012

Eksamenssted: Dragvoll

Eksamenstid: 5 timer

Studiepoeng: 15

Tillatte hjelpemidler: Flg formelsamling: Knut Sydsæter, Arne Strøm og Peter Berck (2006): Matematisk formelsamling for økonomer, 4utg. Gyldendal akademiske. Knut Sydsæter, Arne Strøm, og Peter Berck (2005): Economists' mathematical manual, Berlin.
Enkel kalkulator Citizen SR-270x el. HP 30S.

Sensur: 28. juni 2012

Eksamen består av 2 oppgaver med delspørsmål som alle skal besvares.

Antall sider bokmål: 2
Antall sider nynorsk: 2
Antall sider engelsk: 2

Vedlegg: 2 tabeller

Merk! Det blir sendt automatisk varsel om sensur på e-post. Du kan se hva som er registrert ved å gå inn på Studentweb. Evt andre telefoner om sensur må rettes til instituttet. Eksamenskontoret vil ikke kunne svare på slike telefoner.

Oppgave 1

a) Betrakt modellen gitt ved:

$$y_i = \beta_0 + \beta_1 x_i + u_i$$

Utled minste kvadraters metode (MKM) estimatoren for parameteren β_1 og drøft under hvilke forutsetninger om restleddet u_i denne estimatoren er forventningsrett.

b) Formuler en enkel paneldatamodell og drøft alternative metoder (eller transformasjoner) som kan benyttes for å estimere de ukjente parametrene i modellen. Drøft fordeler og ulemper ved bruk av fixed effects modellen (within groups transformasjonen) og drøft spesielt i hvilken grad denne metoden løser et problem med utelatte eller uobserverbare variable.

c) Du er bedt om å undersøke empirisk om en undervisningsreform gir bedre elevprestasjoner. Du har data for elevprestasjoner i n skoler som har gjennomført reformen og m skoler som ikke har gjennomført reformen. For begge skolegruppene har du data for en periode før og en periode etter reformen. Forklar hvordan du vil gå fram for å teste om undervisningsreformen medførte bedre elevprestasjoner.

Oppgave 2

I en empirisk undersøkelse av lønnsdanning benyttes kvartalsdata for perioden 1979, 2. kvartal til og med 2009, 3 kvartal, i alt 122 observasjoner. I analysen estimeres først følgende modell

$$(1)\quad \ln w_t = \alpha \ln w_{t-1} + \beta_1 \ln p_t + \beta_2 \ln p_{t-1} + \beta_3 \ln z_t + \beta_4 \ln z_{t-1} + \beta_5 \ln ur_t + \beta_6 \ln ur_{t-1} + \beta_0 + v_t$$

der *w* er nominell lønn, *p* er pris, *z* er produktivitet, *ur* er arbeidsledighetsraten i prosent av arbeidsstyrken mens *v* er et stokastisk restledd. Estimering av ligning (1) ved bruk av MKM ga følgende resultater:

$$(2)\quad \begin{aligned} \ln w_t = {} & \underset{(0.08)}{0.53} \ln w_{t-1} + \underset{(0.46)}{0.52} \ln p_t - \underset{(0.47)}{0.05} \ln p_{t-1} + \underset{(0.11)}{0.66} \ln z_t - \underset{(0.12)}{0.18} \ln z_{t-1} - \underset{(0.03)}{0.04} \ln ur_t - \underset{(0.03)}{0.01} \ln ur_{t-1} \\ & - \underset{(0.31)}{0.31}, \qquad SSR = 0.1122 \end{aligned}$$

der tall i parenteser under estimerte parametre er estimerte standardavvik og SSR er summen av kvadrerte avvik.

Merk! Det blir sendt automatisk varsel om sensur på e-post. Du kan se hva som er registrert ved å gå inn på Studentweb. Evt andre telefoner om sensur må rettes til instituttet. Eksamenskontoret vil ikke kunne svare på slike telefoner.

Resultater basert på en forenklet modell er videre gitt ved:

(3) $\ln w_t = \underset{(0.07)}{0.47}\ln w_{t-1} + \underset{(0.07)}{0.51}\ln p_t + \underset{(0.08)}{0.55}\ln z_t - \underset{(0.01)}{0.05}\ln ur_t - \underset{(0.29)}{0.42},\qquad SSR = 0.1147$

a) Forklar hvordan du kan teste hypotesen $\beta_2 = \beta_4 = \beta_6 = 0$ og gjennomfør testen ved bruk av opplysningene gitt i forbindelse med ligning (2) og (3).

b) Test hypotesen $\beta_5 = 0$ ved bruk av resultatene i ligning (2). Test deretter samme hypotese ved bruk av resultatene i ligning (3). Hvorfor blir konklusjonen forskjellig?

c) Forklar hvordan du kan teste hypotesen $\beta_3 = \beta_1$ (lik effekt av pris og produktivitet) og gjennomfør testen ved bruk av resultatene i ligning (4):

(4) $\ln w_t = \underset{(0.07)}{0.47}\ln w_{t-1} - \underset{(0.08)}{0.04}\ln p_t + \underset{(0.08)}{0.55}\left(\ln z_t + \ln p_t\right) - \underset{(0.01)}{0.05}\ln ur_t - \underset{(0.29)}{0.42},\qquad SSR = 0.1147$

d) Drøft kort hva det betyr at estimert verdi på parameteren α er lik 0.47 i ligning (4). Finn deretter den kortsiktige og langsiktige effekten på lønn av (i) en økning i arbeidsledighetsraten med 1 prosent og (ii) en økning i ledighetsraten fra 3% til 4%.

e) En innvending mot å estimere lønnsligningen ved bruk av MKM er at både lønn og pris er endogent bestemt innenfor en simultan modell. Drøft hva slags økonometriske problem dette skaper. Diskuter videre hvordan du vil gå fram for å ta hensyn til dette problemet.

Oppgåve 1

a) Betrakt modellen gitt ved:

$$y_i = \beta_0 + \beta_1 x_i + u_i$$

Utlei minste kvadraters metode (MKM) estimatoren for parameteren β_1 og drøft under kva føresetnader om restleddet u_i denne estimatoren er forventningsrett.

b) Formuler ein enkel paneldatamodell og drøft alternative metodar (eller transformasjonar) som kan nyttes for å estimere dei ukjende parametrane i modellen. Drøft føremoner og ulemper ved bruk av fixed effects modellen (within groups transformasjonen) og drøft særskilt i kva grad denne metoden løyser eit problem med utelatne eller uobserverbare variablar.

c) Du er bedt om å undersøke empirisk om ei undervisningsreform gir betre elevprestasjonar. Du har data for elevprestasjonar i n skular som har gjennomført reforma og m skular som ikkje har gjennomført reforma. For begge skulegruppene har du data for ein periode før og ein periode etter reforma. Forklar korleis du vil gå fram for å teste om undervisningsreforma medførte betre elevprestasjonar.

Oppgåve 2

I ei empirisk undersøking av lønnsdanning nyttes kvartalsdata for perioden 1979, 2. kvartal til og med 2009, 3 kvartal, i alt 122 observasjonar. I undersøkinga vert først fylgjande modell estimert:

$$(1)\quad \ln w_t = \alpha \ln w_{t-1} + \beta_1 \ln p_t + \beta_2 \ln p_{t-1} + \beta_3 \ln z_t + \beta_4 \ln z_{t-1} + \beta_5 \ln ur_t + \beta_6 \ln ur_{t-1} + \beta_0 + v_t$$

der *w* er nominell lønn, *p* er pris, *z* er produktivitet, *ur* er arbeidsløyseraten i prosent av arbeidsstyrken mens *v* er et stokastisk restledd. Estimering av likning (1) ved bruk av MKM ga fylgjande resultat:

$$(2)\quad \begin{aligned} \ln w_t = &\underset{(0.08)}{0.53}\ln w_{t-1} + \underset{(0.46)}{0.52}\ln p_t - \underset{(0.47)}{0.05}\ln p_{t-1} + \underset{(0.11)}{0.66}\ln z_t - \underset{(0.12)}{0.18}\ln z_{t-1} - \underset{(0.03)}{0.04}\ln ur_t - \underset{(0.03)}{0.01}\ln ur_{t-1} \\ &- \underset{(0.31)}{0.31}, \qquad SSR = 0.1122 \end{aligned}$$

der tal i parentesar under dei estimerte parametrane er estimerte standardavvik og SSR er summen av kvadrerte avvik.

Merk! Det blir sendt automatisk varsel om sensur på e-post. Du kan se hva som er registrert ved å gå inn på Studentweb. Evt andre telefoner om sensur må rettes til instituttet. Eksamenskontoret vil ikke kunne svare på slike telefoner.

Resultat basert på ein forenkla modell er vidare gitt ved:

(3) $\ln w_t = \underset{(0.07)}{0.47} \ln w_{t-1} + \underset{(0.07)}{0.51} \ln p_t + \underset{(0.08)}{0.55} \ln z_t - \underset{(0.01)}{0.05} \ln ur_t - \underset{(0.29)}{0.42}$, $\quad SSR = 0.1147$

a) Forklar korleis du kan teste hypotesen $\beta_2 = \beta_4 = \beta_6 = 0$ og gjennomfør testen ved bruk av opplysningane gitt i samband med likning (2) og (3).

b) Test hypotesen $\beta_5 = 0$ ved bruk av resultata i likning (2). Test deretter same hypotese ved bruk av resultata i likning (3). Kvifor vert konklusjonen forskjellig?

c) Forklar korleis du kan teste hypotesen $\beta_3 = \beta_1$ (lik effekt av pris og produktivitet) og gjennomfør testen ved bruk av resultata i likning (4):

(4) $\ln w_t = \underset{(0.07)}{0.47} \ln w_{t-1} - \underset{(0.08)}{0.04} \ln p_t + \underset{(0.08)}{0.55} (\ln z_t + \ln p_t) - \underset{(0.01)}{0.05} \ln ur_t - \underset{(0.29)}{0.42}$, $\quad SSR = 0.1147$

d) Drøft kort kva det betyr at estimert verdi på parameteren α er lik 0.47 i likning (4). Finn deretter den kortsiktige og langsiktige effekten på lønn av (i) ein auke i arbeidsløyseraten med 1 prosent og (ii) ein auke i arbeidsløyseraten frå 3% til 4%.

e) Ei innvending mot å estimere lønnslikninga ved bruk av MKM er at både lønn og pris er endogent bestemt innanfor ein simultan modell. Drøft kva slag økonometriske problem dette skaper. Diskuter vidare korleis du vil gå fram for å ta omsyn til dette problemet.

Question 1

a) Consider the model given by:

$$y_i = \beta_0 + \beta_1 x_i + u_i$$

Derive the ordinary least squares (OLS) estimator for the parameter β_1 and discuss under which assumptions about the error term u_i this estimator is unbiased.

b) Formulate a simple panel data model and discuss alternative methods (or transformations) that can be used to estimate the unknown parameters in the model. Discuss advantages and disadvantages with the use of the fixed effects model (or the within groups transformation) and discuss in particular to what extent this method solves a problem with excluded or unobservable variables.

c) You are asked to investigate empirically whether or not a teaching reform improves student performance. You have data for student performance in n schools that have implemented the reform and m schools that have not implemented the reform. For both groups of schools you have data for a period before and a period after the reform. Explain how you would proceed to test whether or not the teaching reform improves student performance.

Question 2

An empirical study of wage formation uses quarterly data for the period 1979, 2nd quarter to 2009, 3rd quarter, totally 122 observations. In the study, the following model is first estimated:

$$(1)\quad \ln w_t = \alpha \ln w_{t-1} + \beta_1 \ln p_t + \beta_2 \ln p_{t-1} + \beta_3 \ln z_t + \beta_4 \ln z_{t-1} + \beta_5 \ln ur_t + \beta_6 \ln ur_{t-1} + \beta_0 + v_t$$

where *w* is nominal wage, *p* is price, *z* is productivity, *ur* is the unemployment rate in per cent of the labour force whereas *v* is a random error term. Estimating equation (1) using OLS gave the following results:

$$(2)\quad \begin{aligned} \ln w_t = {} & \underset{(0.08)}{0.53} \ln w_{t-1} + \underset{(0.46)}{0.52} \ln p_t - \underset{(0.47)}{0.05} \ln p_{t-1} + \underset{(0.11)}{0.66} \ln z_t - \underset{(0.12)}{0.18} \ln z_{t-1} - \underset{(0.03)}{0.04} \ln ur_t - \underset{(0.03)}{0.01} \ln ur_{t-1} \\ & - \underset{(0.31)}{0.31}, \qquad SSR = 0.1122 \end{aligned}$$

where numbers in parentheses below the estimated parameters are estimated standard errors and SSR is the sum of squared residuals.

Results based on a simplified model are further given by:

$$(3)\quad \ln w_t = \underset{(0.07)}{0.47}\ln w_{t-1} + \underset{(0.07)}{0.51}\ln p_t + \underset{(0.08)}{0.55}\ln z_t - \underset{(0.01)}{0.05}\ln ur_t - \underset{(0.29)}{0.42}, \qquad SSR = 0.1147$$

a) Explain how you can test the hypothesis $\beta_2 = \beta_4 = \beta_6 = 0$ and carry out the test using the information given in relation to equation (2) and (3).

b) Test the hypothesis $\beta_5 = 0$ using the results in equation (2). Test thereafter the same hypothesis using the results in equation (3). Why is the conclusion different?

c) Explain how you can test the hypothesis $\beta_3 = \beta_1$ (equal effect of price and productivity) and carry out the test using the results in equation (4):

$$(4)\quad \ln w_t = \underset{(0.07)}{0.47}\ln w_{t-1} - \underset{(0.08)}{0.04}\ln p_t + \underset{(0.08)}{0.55}\left(\ln z_t + \ln p_t\right) - \underset{(0.01)}{0.05}\ln ur_t - \underset{(0.29)}{0.42}, \qquad SSR = 0.1147$$

d) Discuss briefly what it means that the estimated value of the parameter α is equal to 0.47 in equation (4). Find thereafter the short run and the long run effect on wages of (i) increasing the unemployment rate by 1 percent and (ii) increasing the unemployment rate from 3% to 4%.

e) An objection to estimating the wage equation using OLS is that both wages and prices are endogenously determined within a simultaneous model. Discuss what kind of econometric problem this creates. Discuss further how you would proceed to take this problem into account.

Kommentar til besvarelse i SØK3001, vår 2012, kandidat 10010

Dette er gjennomgående en meget god og velstrukturert besvarelse der alle delspørsmål er klart tilfredsstillende besvart. Kandidatens svar på Oppgave 1b), 1c), 2b) og 2e) trekker opp. På 2b) trekker kandidaten fram at ulik konklusjon mest sannsynlig skyldes multikollinaritet mellom u(t) og u(t-1) i den mest generelle modellen. 2e) honoreres fordi framstillingen er klart relevant for problemstillingen i oppgaveteksten (her hadde en del kandidater et noe liberalt forhold til oppgaveteksten).

Framstillingen er stort sett presis, men på 1a) kunne oppramsingen av alle forutsetningene med fordel vært utelatt. Slik oppgaveteksten er formulert er den eneste forutsetningen av betydning for forventningsretthet at $E(u|X) = 0$.

På oppgave 2d) del ii) må effektene multipliseres med 100 for å få prosentvis lønnsendring. Korttidseffekten av økt ledighet fra 3-4% er tilnærmet -1.4%, langtidseffekten tilnærmet -2.7%.

Kåre Johansen

Oppgave 1

a)

Modellen vi betrakter er:

$$y_i = B_0 + B_1 X_i + u_i$$

Denne modellen kan løses ved minste kvadraters metode. Det vi ønsker å finne er estimert verdi av B_1 altså $\hat{B}_1$. MKM-estimering går ut på at vi ønsker å minimere SSR (sum of square residuale.

Jeg vil nå starte denne oppgaven ved å gjøre rede for en del forutsetninger som må være oppfylt for at $\hat{B}_1$ skal være forventningsrett (og også konsistent.

Som sagt er dette en enkel regresjonsmodell, med en forklaringsvariabel x_i. y_i er den avhengige variablen og vi ønsker å estimere hvilken effekt x_i har på denne. B_0 kan betraktes som konstantleddet (skjæringspunktet med y-aksen. u_i representerer restleddet som fanger opp uobserverbare effekter. Det er også viktig å merke seg at hvis vi har utelatt relevante forklaringsvariabler fra modellen vil u_i også inneholde disse.

Forutsetningene for bruk av MKM-estimering er følgende

MLR.1 Populasjonsmodellen er lineær i sine ~~[illegible]~~ parametere

$$y = B_0 + B_1 x_1 + u$$

MLR.2 Uavhengige utvalg (tilfeldig utvalg)

$Cov(u_i, u_j) = 0 \quad i \neq j$

Dette gir oss da: $y_i = B_0 + B_1 x_i + u_i$

MLR.3 Ingen perfekt multikolinaritet
Med dette så menes det at ~~paran~~ forklaringsvariablene ikke kan skrives som en eksakt lineær kombinasjon av andre forklaringsvariabler

MLR4. No conditonal mean
Men dette så menes det at
$E(u_i | x_i) = 0$ som impliserer at

i) $E(u_i) = 0$ ii) $Cov(u_i, x_i) = 0$

Denne antagelsen kan beskrives som eksogenitetsbetingelsen. Vi har altså at forventningen til restleddet betinget på x_i (som er den eneste forklaringsvariabelen i vår modell) skal være lik 0. Dette impliser altså at kovariansen mellom restleddet og forklaringsvariabelen er lik 0 (noe som igjen betyr at de er uavhengige av hverandre. Når denne forutsetningen om restleddet er oppfylt vil vi ha forventningsrette estimatorer. Men forventningsrette estimatorer mener vi at ~~B1~~ $E(\hat{B}_1) = B_1$.
Hvis ~~denne~~ disse betingelsene i MLR ikke er oppfylt vil vi da ~~ikke~~ ha et endogenitetsproblem i x_i som gjør at vi ikke vil få forventningsrette estimatorer.

MLR.5 Homoskedatisitet
$Var(u_i | x_i) = \sigma^2$
Vi har altså en konstant varianse som er uavhengig av i.

MLR.6 Normalfordelt restledd $u_i \sim N(0, \sigma^2)$
Restleddet normalfordelt med forventning lik 0 og

varianse lik σ^2 (Merk imidlertid at hvis n er stor kan vi støtte oss på sentralgrenseteoremet og denne antagelsen ~~kan~~ trenger da ikke å gjelde "eksakt").

Har nå vist forutsetningene som må være til stede under en MKM-estimasjon av modellen. Det er viktig å merke seg at i vår modell har vi tversnittdata. Videre har vi som sagt kun en forklaringsvariabelen, men forutsetningene vil vært akkurat de samme for en modell med k variabler ($x_{i1}, \ldots x_{ik}$). MLR 4 som er forutsetningen vi gjør om restleddet ville da blitt $E(u_i | X) = 0$ hvor X da er en vektor bestående av $x_{i1}, \ldots, x_{ik}$. Prinsippet vil altså være akkurat det samme bare at u_i nå er betinget på alle forklaringsvariablene.

Har at MLR 1 til MLR 4 vil gi forventningsrette estimatorer hvis oppfylt. Altså at $E(\hat{B}_1) = B_1$
Videre har vi at MLR 1 til MLR 5 vil gi såkalte BLUE estimatorer. Dette vil si at vi har effisiente estimatorer, som er den beste lineære forventningsrette estimatorer med minst varianse.

Har nå gjort rede for forutsetningene for bruk av MKM, og kan dermed finne estimatoren for parameteren B_1.

Har da at modellen er gitt ved: $y_i = B_0 + B_1 x_i + u_i$

Videre kan vi vise at predikert verdi av modellen er gitt ved:
$\hat{y}_i = \hat{B}_0 + \hat{B}_1 x_i$

NTNU

Kandidat nr./*Candidate no.* 10010

Dato/*Date:* 7/6.12 Side/*Page:* 4

Emnekode/*Subject* søk 3001 Antall ark/*Number of pages:* 36

Denne kolonnen er forbeholdt sensor

This column is for external examiner

Vi definerer så residualene $\hat{u}_i$ som er

$y_i - \hat{y}_i = \hat{u}_i$

Dette gir oss da at

$\hat{u}_i = y_i - \hat{B}_0 - \hat{B}_1 x_i$

Vi kan tenke på residualene som "motstykket" til restleddet u_i".

~~[illegible]~~

MkM går da ut på at vi vil minimere $SSR = \sum_{i=1}^{n}(\hat{u}_i)^2$

Dette gir oss følgende minimeringsproblem:

$$\text{Min} \sum_{i=1}^{n}(\hat{u}_i)^2 = \sum_{i=1}^{n}(y_i - \hat{B}_0 - \hat{B}_1 x_i)^2 \quad \text{mhp } \hat{B}_1 \text{ og } \hat{B}_0$$

(vil videre ikke bruke notasjon på summetegnet av enkelhets grunn).

FOB:

$$\frac{\partial SSR}{\partial \hat{B}_0} = 2\sum(y_i - \hat{B}_0 - \hat{B}_1 x_i)(-1) = 0 \qquad 1)$$

$$\frac{\partial SSR}{\partial \hat{B}_1} = 2\sum(y_i - \hat{B}_0 - \hat{B}_1 x_i)(-x_i) = 0 \qquad 2)$$

Fra 1) finner vi da at:

$$\sum(y_i - \hat{B}_0 - \hat{B}_1 x_i) = 0$$

$$\sum y_i - n\hat{B}_0 - \hat{B}_1 \sum x_i = 0$$

$$\rightarrow \hat{B}_0 = \bar{y} - \hat{B}_1 \bar{x}$$

Setter så inn igjen for dette utrykket i minimeringsproblemet, altså erstatter $\hat{B}_0$ her. Dette gir følgende problem:

$$\text{Min} \sum(y_i - \bar{y} - \hat{B}_1(x_i - \bar{x}))^2 \quad \text{mhp } \hat{B}_1$$

Den gule kopien beholder du/*Please keep the yellow page*

Tapir Uttrykk 05.2012

FOB:

$$\frac{\partial SSR}{\partial \hat{B_1}} = 2\Sigma(y_i - \bar{y} - \hat{B_1}(x_i - \bar{x}))(-1)(x_i - \bar{x}) = 0$$

som kan ~~.....~~ omformes og skrives på følgende måte:

$$\Sigma(y_i - \bar{y})(x_i - \bar{x}) - \hat{B_1}\Sigma(x_i - \bar{x})(x_i - \bar{x})$$

Dette vil gi oss følgende estimator for B_1

$$\hat{B_1} = \frac{\sum_{i=1}^{n}(y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

✓

Har nå funnet estimatoren $\hat{B_1}$ for B_1, og har tidligere i oppgaven vist hvilke forutsetninger om u_i som må være oppfylt for at estimatoren skal være forventningsrett

For å vise at $\hat{B_1}$ faktisk er forventningrett har vi at:

$$\bar{y} = B_0 + B_1\bar{x} + \bar{u}$$

Hvis vi trekker dette fra y_i har vi at

$$y_i - \bar{y} = B_1(x_i - \bar{x}) + u_i - \bar{u}$$

Setter så dette inn i uttrykket for $\hat{B_1}$:

$$\hat{B_1} = \frac{\Sigma(B_1(x_i - \bar{x}) + u_i - \bar{u})(x_i - \bar{x})}{\Sigma(x_i - \bar{x})^2}$$

$$\hat{B_1} = B_1 + \frac{\Sigma(u_i - \bar{u})(x_i - \bar{x})}{\Sigma(x_i - \bar{x})^2}$$

$$\hat{B}_1 = B_1 + \frac{\sum u_i (x_i - \bar{x})}{\sum (x_i - \bar{x})^2}$$

siden vi har at

$$-\frac{\bar{u}\sum(x_i-\bar{x})}{\sum(x_i-\bar{x})^2} = 0$$

Hvis vi da tar forventningen av dette utrykket har vi at

$$E(\hat{B}_1 | x_i) = B_1 + \frac{\sum(x_i-\bar{x})E(u_i|x_i)}{\sum(x_i-\bar{x})^2}$$

(Vi tar forventningen betinget på x. Har fra forutsetningene at $E(u_i|x_i)=0$, slik at det vi stor igjen med er at

V

$$E(\hat{B}_1 | x_i) = B_1$$

altså har vi at estimatoren er forventingsrett

Tilsvarende kan vi vise at den også vi være konsistent
Tar da utgangspkt i.

$$\hat{B}_1 = B_1 + \frac{\frac{1}{n}\sum u_i(x_i-\bar{x})}{\frac{1}{n}\sum(x_i-\bar{x})^2}$$

Tar så sannsynlighetsgrensen og antar at de empiriske momentene konvergerer mot sine respektive teoretiske moment

$$\text{plim}(\hat{B}_1) = B_1 + \frac{\text{plim}(\frac{1}{n}\sum u_i(x_i-\bar{x}))}{\text{plim}(\frac{1}{n}\sum(x_i-\bar{x})^2)}$$

$$\text{plim}(\hat{B}_1) = B_1 + \frac{\text{Cov}(u_i, x_i)}{\text{Var}(x_i)}$$

Som vi ser har vi under MLR 4 antatt at $\text{Cov}(u_i, x_i)=0$ slik at estimatoren $\hat{B}_1$ også vil være konsistent.

NTNU

Kandidat nr./*Candidate no.* 10010

Dato/*Date:* 7/6.12 Side/*Page:* 7

Emnekode/*Subject* 10010 Antall ark/*Number of pages:* 36

Denne kolonnen er forbeholdt sensor
This column is for external examiner

V

Har nå utleddet estimatoren til β_1 og vist at denne både vil være forventningrett og konsistent når antagelsene vi har gjort om restleddet holder. Det er altså viktig at forklaringsvariabelen ikke er korrelert med restleddet, og at forventningen til restleddet er lik 0. Hvis disse betingelsene ikke er oppfylt vil vi få forventningsskjeve og inkonsistente estimatorer.

Den gule kopien beholder du/*Please keep the yellow page*

Tapir Uttrykk 05.2012

Oppgave 1 b):

Har nå paneldata som betyr at vi har observasjoner om samme i på ulike tidspunkt. Dette kan f.eks være at vi følger en gruppe husholdninger i så så mange år. Hovedessensen i paneldata er at det er de samme i vi følger. Dette fører til at vi har to typer variasjon som vi kaller between og within. Between variasjon kan ses på som variasjon mellom individer (mellom i), mens within variasjon kan ses på som variasjon for i ~~over~~ tidspunkter. Merk at en balansert paneldatasett betyr at vi har observasjoner om alle i på alle tidspkt (NT)

En enkel paneldatamodell kan framstilles på følgende måte:

$$y_{it} = \cancel{\alpha} + B_1 x_{it1} + \gamma_1 Z_{i1} + \delta_0 Z^*_{t1} + u_{it}$$

Ser her at vi har 3 typer variabler:

i) Variabler som varierer både between og within (altså fotnote it)

ii) Variabler som kun varierer mellom individ, men er konstante over tid (Z_{i1})

iii) Felles aggregerte variabler, som altså er felles for alle i men vil variere over tid (Z^*_{ti})

(Dette kan være effekter som inflasjon, nasjonall lønnsoppgjør osv).

Alternativ til de felles aggregerte effektene (Z^*_{ti}) har vi at vi kan innføre tidsdummyer for (T-1) av periodene. Merk at hvis vi innfører slike tidsdummyer kan vi ikke inkludere slike aggregerte variabler da dette vil gi oss multikolinoritetsproblemet. Vi har at innføring av tidsdummyer renske ut all feles variasjon over tid, og kan være en

Denne kolonnen er forbeholdt sensor
This column is for external examiner

Kur ting å innføre i en paneldata modell hvis vi er bekymret for at vi har slike aggregerte effekter som er fanget opp i u_{it}

Vi har altså at vi definerer en dummyvariabel for (T-1) av periodene:

$$D2_t = \begin{cases} 1 & \text{hvis } t=2 \\ 0 & \text{ellers} \end{cases} \qquad DT_t = \begin{cases} 1 & \text{hvis } t=T \\ 0 & \text{ellers} \end{cases}$$

Dette vil da gi oss modellen:

$$y_{it} = \alpha + B_1 x_{it1} + \gamma_1 z_{i1} + \delta_2 D2_t + \ldots + \delta_T DT_t + u_{it}$$

Alternativt til den individspesifikke (tverrsnittsspesifikke) komponenten har vi at vi kan innføre dummyvariabler for N-1 av tverrsnitt enhetene. Dette vil da si at vi ikke kan inkludere slike ~~tverrsnittsspesifikke~~ komponenter da dette vil gi et multikolinaritetsproblem. Her at innføringen av slike ~~...~~ tversnittsdummyer vil renske bort all variasjon mellom individ. Dette vil være ekvivalent med å innføre et individspesifikt konstantledd α_i. Modellen vil da være

$$y_{it} = \alpha_i + B_1 x_{it1} + \delta_2 D2_t + \ldots + \delta_T DT_t + u_{it}$$

Ser nå videre på hvilke forutsetninger som må være oppfylt ved estimering av paneldata. Velger da forenkelhetsskyld å se bort ifra felles aggregerte effekter (men merk at modellen kan ~~fosco~~ ved å inkludere dummyvariabler for disse)

Antar nå at vi har modellen

$$y_{it} = \alpha + B_1 x_{it1} + \gamma_1 z_{it} + u_{it}$$

Vi ~~har~~ ~~følgende forutsetninger~~

Denne kolonnen er forbeholdt sensor
This column is for external examiner

Videre antar vi at vi kan gjøre en dekomponering av restleddet u_{it}. Har at $u_{it} = \eta_i + \varepsilon_{it}$

η_i kan ses på som individspesifikke effekter. Varierer mellom individ men ikke over tid. Kan feks være aldersammensetning, demokratiske forhold, historiske momenter osv

ε_{it} - ideosynkratisk komponent. Som varier både mellom i og over tid (fanger også opp uobserverte effekter

Har da følgende forutsetninger

i) $E(\varepsilon_{it} | X) = 0$ (antar X som en vektor bestående av alle forklaringsvariabler. (vårt tilfelle z_{it} og x_{it})

Dette er eksogenitetsbetingelsen som må være oppfylt for å gi forventningsrette estimatorer

ii) $E(\varepsilon_{it}, \varepsilon_{js}) = \begin{cases} \sigma_\varepsilon^2 & \text{hvis } i=j \text{ og } s=t \\ 0 & \text{ellers} \end{cases}$

Denne betingelsen viser at ε_{it} vil ha konstant varians og ingen seriekorrelasjon

iii) $E(\eta_i \eta_j) = \begin{cases} \sigma_\eta^2 & \text{hvis } i=j \\ 0 & \text{ellers} \end{cases}$

iv) $E(\varepsilon_{it} \eta_j) = 0$ for alle i, t, j. Som viser at de to komponentene er uavhengige av hverandre

v) $E(\eta_i | X) = 0$ Det er ofte denne eksogenitetsbetingelsen det er knyttet usikkerhet til ved paneldata. For å kunne bruke pooled MKM estimering er vi avhengig av at denne betingelsen er oppfylt for å få forventningsrette estimatorer. Hvis vi har at $E(\eta_i | X) \neq 0$

vil pooled MKM gi forventningsskjeve og inkonsitente estimatorer.

Jeg har nå formulert en paneldatamodell og vist forutsetninger om restleddskomponentene som må være oppfylt. Hvilken estimeringsmetode vi velger å bruke for estimering av vår paneldatamodellen avhenger av betingelse v). Hvis vi har at den er oppfylt kan vi bruke pooled OLS. Alternativt kan vi som beskrevet tidligere benytte tidsdummye og tverrsnittsspesifikke dummyer hvis vi har mistanke om at vi har individspesifikke eller aggregerte telles faktoer som er fanget i restleddet (merk at vi da ikke kan inkluder Z_i og Z_t variabler), og deretter bruke pooled ~~OLS~~ MKM.

Pooled MKM går ut på følgende:

Vi definerer det globale gjennomsnitt av panel data modellen vår:

$$y_{it} = \alpha + B_1 X_{it1} + \gamma_1 Z_{i1} + u_{it}$$

$$\bar{y} = \frac{1}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N} y_{it} \qquad \bar{u} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} u_{it}$$

$$\bar{X}_1 = \frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n} X_{it} \qquad \bar{Z}_1 = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} Z_{i1} = \frac{T\sum_{i=1}^{N} Z_{i1}}{NT}$$

Dette gir oss da følgende likning:

$$\bar{y} = \alpha + B_1 \bar{X}_1 + \gamma \bar{Z}_1 + \bar{u}$$

Vi vil nå regne variablene som avvik fra sitt globale gjennomsnitt:

$$y_{it} - \bar{y} = B_1(X_{it1} - \bar{X}_1) + \gamma(Z_{i1} - \bar{Z}_1) + u_{it} - \bar{u}$$

√ Kan nå bruke MKM på denne likningen.

Det vi ser med pooled ~~OLS~~ MKM er at vi benytter oss av all variasjonen i modellen, både between og within. I tillegg har vi at vi ved denne estimeringsmetoden kan estimere

Data √

effekten av individspesifikk (tverrsnittsspesifikke) komponenter.
Det som imidlertid ofte er en ulempe med denne ~~modell~~ estimeringen er at vi ofte ikke med sikkerhet kan si at $E(n_i|t) = 0$. Hvordan kan estimere parameterne i en paneldata modell hvis vi har at denne ikke er oppfylt ($E(n_i|t) = 0$)
Vi antar videre at vi innfører et individspesifikt konstantledd $\alpha_i = \alpha + \eta_i$.
Slik at modellen vi nå har er:

$$y_{it} = \alpha_i + B_1 x_{it1} + \gamma_1 z_{i1} + u_{it}$$

~~Sammenhengen~~
Vi kan da bruke Fixed effects metoden (FE) eller first-difference metoden) for å estimere parameterne i denne modellen. Det vi vil vise er at vi kan transformere modellen slik at vi kan kvitte oss med dette endogenitetsproblemet vi står ovenfor.
Ser da først Fixed effekts modellen (også kalt within groups transformasjon. (Merk at dette vil være ekvivalent med å innføre tverrsnittsdummyer).
Har da at vi starter med å ta gjennomsnittet over tid:

$$\bar{y}_i = \frac{1}{T}\sum_{t=1}^{T} y_{it} \qquad \bar{\alpha}_i = \frac{1}{T}\sum_{t=1}^{T}\alpha_i = \frac{T\alpha_i}{T}$$

$$\bar{x}_{i1} = \frac{1}{T}\sum_{t=1}^{T} x_{it} \qquad \bar{z}_{i1} = \frac{1}{T}\sum_{t=1}^{T} z_{i1} = \frac{T z_{i1}}{T}$$

$$\bar{\varepsilon}_i = \frac{1}{T}\sum_{t=1}^{T}\varepsilon_{it}$$

Dette gir oss følgende likning:

$$\bar{y}_i = \alpha_i + B_1\bar{x}_{i1} + \gamma_1 z_{i1} + \bar{\varepsilon}_i$$

Regner så variablene som avik fra dette individspesifikke gjennomsnittet (gjennomsnittet over tid)

$$y_{it} - \bar{y}_i = B_1(x_{it1} - \bar{x}_{i1}) + \varepsilon_{it} - \bar{\varepsilon}_i$$

$$\ddot{y}_{it} = B_1 \ddot{x}_{it} + \ddot{\varepsilon}_{it}$$

~~Kan~~ Kan så estimere dette ved bruk av MKM. Det som er viktig å legge merke til her er at vi har transformert bort vårt problem med at vi har uobserverbar heterogenitet (representert ved n_i) eller et utelatt variabel problem (også representert ved n_i) slik at vi ved estimering av likningen ovenfor vil få forventningsrette estimatorer. Fordelen ved bruk av within groups transformasjon er at vi har høyde for at vi har et endogenitetsproblem og korrigerer for dette. Dette gjør at vi kan få forventningsrette estimatorer. Ulemper med denne type transformasjon er at den benytter mindre variasjon enn ved pooled OLS, da vi ved within groups transformasjon renser ut all variasjon mellom individ, og utnytter oss kun av tidsvariasjon (within variasjon). Det at vi utnytter mindre variasjon kan fører ~~AAA~~ til at standardfeil kan øke sammenlignet med pooled OLS. En annen ulempe ved å bruke within groups-transformasjon er at vi mister muligheten til å estimere effekten av eventuelle individspesifikke / tverrsnittspesifikke komponenter da disse også transformeres bort fra modellen.

Et annet alternativ til bruk av Fixed effects modeller er first-differencing. FD gir følgende resultat

$$\Delta y_{it} = B_1 \Delta x_{it1} + \Delta \varepsilon_{it}$$

hvor vi har at

$$\Delta y_{it} = y_{it} - y_{it-1} \qquad \Delta \varepsilon_{it} = \varepsilon_{it} - \varepsilon_{it-1}$$

$$\Delta x_{it1} = x_{it1} - x_{it-1,1}$$

NTNU

Kandidat nr./Candidate no. 10010

Dato/Date: 7/6.12 Side/Page: 14

Emnekode/Subject SØK 3001 Antall ark/Number of pages: 36

Denne kolonnen er forbeholdt sensor
This column is for external examiner

Ser her at vi ved FD vil ~~[illegible]~~ kunne transformere bort eventuelle endogenitetsproblem. På samme måte som ved FE utnytter vi kun variasjon over tid, og kan ikke estimere effekten av eventuelle individspesifikke komponenter.
Merk at vi vil ha forskjellig restledd ved FE og FD.

Har nå vist hvordan vi på ulike måter kan estimere parameterne i en paneldatamodellen. Hvilken estimeringsmetode vi bruker avhenger av betingelsen $E(\eta_i|X)$. Hvis denne er oppfylt kan vi bruke pooled MKM som vil gi forventningsrette estimatorer. Alternativt kan vi innføre tidsdummyer og/eller tverrsnittsspesifikke dummyer hvis vi har mistanke om at vi har ~~[illegible]~~ henholdsvis felles aggregerte variabler eller individspesifikke variabler som påvirker modellen men ikke er inkludert som forklaringsvariabler (alternativt er uobserverbare)
Har sett at hvis vi har uobserverbare variabler kan dette representeres med η_i (som også ~~[illegible]~~ vil inkludere utelatte variabler). Da har vi at $E(\eta_i|X) \neq 0$ og vi vet da at pooled MKM vil gi forventningsskjeve estimatorer. En alternativ estimeringsmetode er da withingroups transformasjon. Ved å utføre dette vil vi transformere bort den uobserverte variabelen (også utelatt variabel problem) og vi vil på denne måten kunne oppnå forventningsrette estimator. Ulemper ved dette er at vi da kun utnytter variasjonen over tid, samt at hvis ~~[illegible]~~ vi ønsker å estimere effekten av en tverrsnittsspesifikk komponent vil ikke dette kunne være mulig
(Merk at vi også kunne ha brukt instrumentvariabel metoden for å løse problemet med utelatt forklaringsvariabel)

ok

Oppgave 1 c)

Skal undersøke empirisk om en undervisningsreform gir bedre elevprestasjoner.
Har data for elevprestasjoner i n skoler som har gjennomført reformen.
Data for m skoler som ikke har gjennomført reformen.
Data for en periode før og en periode etter reformen
Skal nå forklare hvordan jeg vil gå fram for å teste om undervisningsreformen medførte bedre elevprestasjoner.

Starter da med å definer 2 typer dummievariabler

V $D2 = \begin{cases} 1 \text{ hvis periode etter reform} \\ 0 \text{ hvis periode før reform} \end{cases}$

$DT = \begin{cases} 1 \text{ hvis gjennomført reform} \\ 0 \text{ hvis ikke gjennomført reform} \end{cases}$ V

Definerer y_{it} som elevprestasjoner. ~~[illegible]~~
Definisjonene ovenfor gir oss følgende likning:

V $$y_{it} = B_0 + B_1 D2 + \delta_0 DT + \delta_1 D2\,DT$$

Forklaring til relasjonen:
Har at hvis perioden før gjennomført reform vil effekten på elevprestasjon være B_0 hvis skole som ikke gjennomført reforme. Hvis skolen gjennomførte reform vil effekten på y_{it} være $B_0 + \delta_0$

I perioden etter reform har vi at effekten på elevprestasjoner er gitt ved $B_0 + B_1$ for skoler som ikke gjennomførte reformen. Mens for skoler som gjennomførte reformen vil effekten være $\beta_0 + \beta_1 + \delta_0 + \delta_1$

For å gjøre ting litt klarer kan vi illustrere dette i et diagram

	FØR	Etter	Etter - før
Kontroll	β_0	$\beta_0 + \beta_1$	β_1
"Behandling"	$\beta_0 + \delta_0$	$\beta_0 + \beta_1 + \delta_0 + \delta_1$	$\beta_1 + \delta_1$
"Behandling" - kontroll	δ_0	$\delta_0 + \delta_1$	δ_1

Merk at vi her ser på kontrollgruppen som den skolen som ikke gjennomførte reformen, mens "behandlingsgruppen" vil være skolen som gjennomførte reformen

Utifra diagrammet ovenfor er det parameteret δ_1 vi ønsker å teste (også kalt difference-in-difference estimatoren).
Det er altså interaksjonsleddet $\delta_1 D2\ DT$ som vil si noe om innføringen av undervisningsreformen medførte bedre elevprestasjoner. Merk at $\delta_0 DT$ viser eventuelle forskjeller som var mellom de to gruppene før innføringen av reformen, mens $\beta_1 D2$ viser at ~~[illegible]~~ forskjellen fra perioden før reform og etter (som vil være uavhengig av reformen).
Som sagt er det δ_1 vi vil ønske å teste. Dette kan vi gjøre ved bruk av en t-test (siden vi kun har en restriksjon)
Det vi da ønsker å teste er om δ_1 er signifikant forskjellig fra 0, altså $H_0: \delta_1 = 0$. Hvis vi kan avvise denne hypotesen kan vi konkludere med at undervisningsreformen medførte bedre elevpresentasjoner. Kommer tilbake til formell beskrivelse av t-test i neste oppgave. (Oppg 2b)).

Alternativhypotesen vil være å teste om $H_A: \delta_1 > 0$

Det vi altså ønsker å teste er:

$$H_0: \delta_1 = 0 \quad vs \quad H_A: \delta_1 > 0$$

Vi har en ensidig t-test

Testatistikken vil vi finne fra:

$$t^{obs} = \frac{\hat{\delta}_1 - \delta_1^*}{se(\hat{\delta}_1)} = \frac{\hat{\delta}_1}{se(\hat{\delta}_1)}$$

t^{crit} vil vi finne fra t-tabellen, avhenger av frihetsgradene vi har og valgt signifikansnivå.

Vi har da at hvis $|t^{obs}| > |t^{crit}|$ vil vi kunne forkaste H_0 noe som ~~[illegible]~~ indikerer at Undervisnings-reformen medførte bedre elevprestasjoner.

(Alternativt kunne vi har brukt en Chow-test, hvor vi tillater for ulikt skjæringspunkt i mellom de to periodene)

Denne kolonnen er forbeholdt sensor

This column is for external examiner

Oppgave 2

Har nå en dynamisk modell (tidsserieanalyse) gitt ved

$$\ln w_t = \alpha \ln w_{t-1} + B_1 \ln p_t + B_2 \ln p_{t-1} + B_3 \ln z_t + B_4 \ln z_{t-1} + B_5 \ln ur_t + B_6 \ln ur_{t-1} + B_0 + v_t$$

w - nomiell lønn

p - pris

z - produktivitet

ur - arbeidsledighetsraten i % av arbeidsstyrken

v - stokastisk restledd

Estimering av likning ved bruk av MKM:

2) $\ln w_t = 0{,}53 \ln w_{t-1} + 0{,}52 \ln p_t - 0{,}05 \ln p_{t-1} + 0{,}66 \ln z_t - 0{,}16 \ln z_{t-1} - 0{,}04 \ln ur_t - 0{,}01 \ln ur_{t-1}$

Resultater av en forenklet modell er gitt ved

3) $\ln w_t = 0{,}47 \ln w_{t-1} + 0{,}31 \ln p_t + 0{,}55 \ln z_t - 0{,}05 \ln ur_t - 0{,}42$

a) Vil teste hypotesen $B_2 = B_4 = B_6 = 0$

Det vi altså vil er å teste $H_0: B_2 = B_4 = B_6 = 0$ vs H_a: felles ≠ 0

Vi ønsker å teste flere restriksjoner og benytter oss derfor av en F-test. Før jeg formelt skal teste H_0 vil jeg gi en forklaring av hva F-testen går ut på. Testobservatoren i en slik F-test er gitt ved

$$F_{obs} = \frac{(SSR_r - SSR_{ur})/q}{SSR_{ur}/(n-k-1)}$$

NTNU

Kandidat nr./Candidate no. 10010

Dato/Date: 7/6.12 Side/Page: 19

Emnekode/Subject SØK 3601 Antall ark/Number of pages: 36

Denne kolonnen er forbeholdt sensor

This column is for external examiner

Har at:

- SSRr : sum of square residuals fra restricted modell (modell pålagt restriksjonene gitt ved null hypotesen)
- SSRur: sum of square residuals fra unrestricted modell (altså den generelle modellen uten noen restriksjoner)
- q er antall restriksjoner, og er frihetsgrader (df i telleren: dfr - dfur)
- (n-k-1) er frihetsgradene i nevneren. Hvor n står for antall observasjon, k står for antall parametere i unrestricted modell (-1 pga konstantleddet).

Det vi altså ønsker å gjøre er å sammenligne den relative endringen i SSR når vi pålegger modellen restriksjonen. Hvis vi har at denne relative økningen er stor kan vi konkludere med at restriksjonene ikke er gyldig. Altså har vi at vi ser på SSRr relativt til SSRur. En F-test tar utgangspunkt i goodness-of fit mesure (kan alternativt også bruke R^2). Vil altså se hvordan forklart variasjon endrer seg når man pålegger den generelle modellen restriksjoner. Hvis denne endringen stor avviser vi hypotesen. Forkastningskriteriet vil avhenge av valgt signifikansnivå og frihetsgradene i telleren og nevneren. Merk at SSRr alltid vil være større enn SSRur. Dette er fordi vi ved MKM vil minimere SSR. Og vi vet at ved å tillegge modellen flere forklaringsvariabler vil forklart variasjon alltid øke.

Grunnen til at vi må bruke en F-test når vi har flere restriksjoner i ønsker å teste er fordi en t-test ikke pålegger restriksjoner på alle andre parameterne i modellen.

Det er viktig å merke seg at ved en F-test tester vi samlet signifikans mot samlet avvisning.

Denne kolonnen er forbeholdt sensor
This column is for external examiner

Har at forkastningskriteriet er gitt ved $|F^{obs}| > |F^{crit}|$
Siden en F-fordeling ikke er symmetrisk vil den være veldig vanskelig å tegne, så velger å ikke gjøre dette.
Hvis H_0 er sann har vi altså at $F^{obs} \sim F_{q,(n-k-1)}$.
Kritiske verdi finner vi i tabellen for F-fordelingen.

Skal nå gjennomføre den formelle testen:

1) Vi former hypotesen: $H_0: B_2 = B_4 = B_6 = 0$
H_a: felles $\neq 0$

2) Finner teststatistikk

$$F^{obs} = \frac{(SSR_r - SSR_{ur})/q}{SSR_{ur}/(n-k-1)}$$

Har at modell 2) vil være unrestricted modell (generell modell)
Modell 3) vil være restricted modell (pålagt restriksjoner)
Antall restriksjoner: $q = 3$
$n = 122$ og $k = 7$ (husk på konstantleddet)

$$F^{obs} = \frac{(0{,}1147 - 0{,}1122)/3}{0{,}1122/(122-7-1)} = 0{,}847$$

3.) Finner så kritiske verdi fra F-tallet. Velger $\alpha = 0{,}05$ (5% signifikansnivå). Frihetsgrader i teller = 3. Frihetsgrader i nevner = 114

Dette gir $F^{crit} = 2{,}68$

4) Sammenligner og konkluderer:
$F^{crit} > F^{obs} \rightarrow 2{,}68 > 0{,}847$

NTNU

Kandidat nr./Candidate no. 10010

Dato/Date: 7/6.12 Side/Page: ~~28~~ 21

Emnekode/Subject SØK 3001 Antall ark/Number of pages: 36

Denne kolonnen er forbeholdt sensor

This column is for external examiner

Dette betyr at vi ~~kan~~ ikke kan forkaste H_0. Dette impliserer da at modell 3) er en gyldig forenkling av modell 2)

Den gule kopien beholder du/Please keep the yellow page

NTNU

Kandidat nr./*Candidate no.* 10010

Dato/*Date:* 7/6.12 Side/*Page:* 22

Emnekode/*Subject* søk 3001 Antall ark/*Number of pages:* 36

Denne kolonnen er forbeholdt sensor

This column is for external examiner

b) Vi ønsker å teste hypotesen $B_3 = 0$ ved likning 2).
Det vi ønsker å tester er altså

$$H_0: B_3 = 0 \quad vs \quad H_A: B_3 \neq 0$$

Vi har altså en to-sidig test. Før jeg tester denne hypotesen formelt vil jeg beskrive utgangspunktet for en t-test. Se neste side.

→

Den gule kopien beholder du/*Please keep the yellow page*

Tapir Uttrykk 05.2012

En t-test for utgangspunkt i MKM forutsetningene som jeg beskrev i oppgave 1 a). Når disse forutsetningene er oppfylt har vi at:

$$Var(\hat{\beta}_j | X) = \frac{\sigma^2}{SST_j(1-R_j^2)}$$

σ^2 - befolkningsrestleddsvarians
$SST_j = \Sigma(x_{ij} - \bar{x}_j)^2$
R_j^2 fra regresjonen av x_j på alle de andre forklaringsvariablene.

Har da at $sd(\hat{\beta}_j) = \frac{\sigma}{\sqrt{SST_j(1-R_j^2)}}$

Som viser standardaviket

Det som imidlertid er er at σ^2 er variansen til restleddet for populasjonen og må derfor estimeres. Vi vil finne en forventningsrett estimator $E(\hat{\sigma}^2) = \sigma$. Har at dette er gitt ved $\hat{\sigma}^2 = \frac{\Sigma(\hat{u}_i)^2}{n-k-1}$ (NB viktig å korrigere for frihetsgrader)

Dette gir oss da estimert varianse:

$$\widehat{Var}(\hat{\beta}_j | X) = \frac{\hat{\sigma}^2}{SST_j(1-R_j^2)}$$

Som videre gir estimert standardavik (standardfeil):

$$se(\hat{\beta}_j) = \frac{\hat{\sigma}}{\sqrt{SST_j(1-R_j^2)}}$$

Under disse forutsetningene har vi da at

$\hat{\beta}_j \sim N(\beta_j, Var(\hat{\beta}_j))$ er normalfordelt med forventning lik β_j og varianse $Var(\hat{\beta}_j)$

$\frac{\hat{\beta}_j - \beta_j}{sd(\hat{\beta}_j)} \sim N(0,1)$ vil være standard normalfordelt med forventning 0 og varianse 1

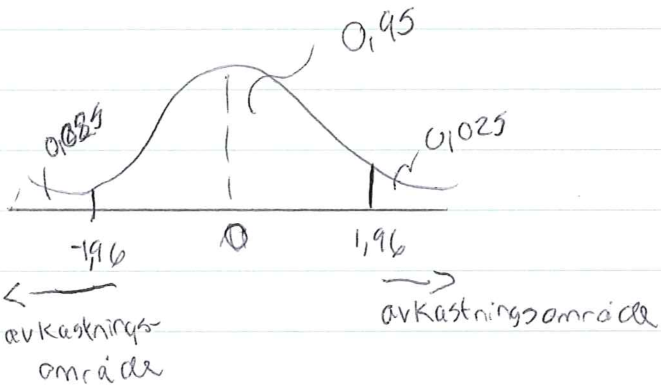
Denne kolonnen er forbeholdt sensor
This column is for external examiner

Dette gir da estimert standardavik

$\frac{\hat{B}_1 - B_1}{se(\hat{B}_1)} \sim t_n$ som vi vær t-fordelt med n frihetsgrader.

Vi ønsker å teste $H_0 : B_i = 0$ mot $H_A : B_i \neq 0$
Dette er en to-sidig test. Hvis vi har at $df > 120$ vil vi og vi antar at H_0 er sann vil vi ha 5% sannsynlighet (ved $\alpha = 0,05$) for å observere verdier større enn 1,96 eller mindre enn −1,96). $t^{obs} = \frac{\hat{B}_i}{se(\hat{B}_i)} \sim n-k-1$

vil være t-fordel med $(n-k-1)$ frihetsgrader (da antall >120)



Ved signifikansnivå 0,05 har vi at det er en 5% sannsynlighet for å forkaste H_0 når den i virkeligheten er sann. Det vi ønsker å teste ved en t-test er hvor mange standardavik $\hat{B}_1$ er fra sin hypotetiske verdi (under H_0), for så å definere en kritiske verdi hvor vi ikke lenger kan anta at estimert og hypotetiske verdi er lik.
Forkastningskriteriet: $|t^{obs}| > |t^{crit}|$. Finner kritiskverdi i tabellen for t-fordeling (avhenger av frihetsgrader og valgt signifikansnivå.
Merk at vi har forskjell på en en-sidig og to-sidig test.
Ensidig vil vil alternativ hypotesen vær $B_i > 0$ eller $B_i < 0$
Ved to sidig test: $H_A: B_i \neq 0$

NTNU

Kandidat nr./*Candidate no.* 10010

Dato/*Date:* 7/6.17 Side/*Page:* 25

Emnekode/*Subject* Sosø 3001 Antall ark/*Number of pages:* 36

Denne kolonnen er forbeholdt sensor

This column is for external examiner

Kan nå gjennomføre den formelle testen:

1) Konstruerer hypotesen:

$H_0: B_5 = 0$ vs $H_a: B_5 \neq 0$

2) Finner teststatistikk

~~TN~~ $t^{obs} = \dfrac{\hat{B}_i - B_i}{se(\hat{B}_i)}$

✓ $t^{obs} = \dfrac{-0{,}04}{0{,}03} = -1{,}33$

3) Finner kritiske verdi: $n = 122$

$k = 7 \quad (k-1 = 8)$

$\alpha = 0{,}05$ Husk at vi har 2sidig test

$t^{crit} = 1{,}98$

4) Sammenligner og konkluderer

$|t^{obs}| < |t^{crit}| \rightarrow |1{,}33| < |1{,}98|$

✓ Vi kan ikke forkaste hypotsen. Det ser da ut til at arbeidsledighetsraten ikke har påvirkning på nominell lønn

NTNU

Kandidat nr./Candidate no. 10010

Dato/Date: 7/6.12 Side/Page: 26

Emnekode/Subject SØK 3001 Antall ark/Number of pages: 36

Denne kolonnen er forbeholdt sensor
This column is for external examiner

skal nå teste samme hypotese ved bruk av likning 3)

1) $H_0: B_3 = 0$ vs $H_A: B_3 \neq 0$

2) $t^{obs} = \frac{-0,05}{0,01} = -5$

3) $t^{crit} = 1,98$ (frihetsgrader (122-5))

4) $|t^{crit}| < |t^{obs}| \rightarrow |1,98| < |5|$

Dette betyr at vi kan forkaste H_0 til fordel for H_A. Det ser ut til at i modell tre vil arbeidsledighetsraten ha innvirkning på nominell lønn.

Ser at vi får forskjellig konklusjon ved test av $B_3 = 0$ ettersom vi bruker resultatene i modell 2) eller modell 3).
Grunnen til dette er at i modell 2) så er lagget verdi av arbeidsledighetsraten inkluder Ur_{t-1}. I modell 3) hvor Ur_{t-1} ikke er inkludert kan vi tenke oss at denne komponentet vil ligge i restleddet. Vi har at Ur_t og Ur_{t-1} vil være korrelert (rimelig å anta at arbeidsledigheten år t avhenger av arbeidsledigheten år t-1). Når begge disse forklaringsvariablene er inkludert som i modell 2) vil det være vanskelig å skille effekten av dem fra hverandre. Vi har altså et multikolinaritetsproblem. Altså vil det i modell 2 være vanskelig å skille de partielle effektene av Ur_t og Ur_{t-1} fra hverandre. Vi har altså at konklusjonen blir forskjellig fordi i likning 3) har vi utelatt U_{t-1} slik at vi ikke vil ha vanskligheter med å skille mellom de partielle effekten av U_t og U_{t-1} som vi vil ha i likning 2, hvor begge er inkludert

NTNU

Kandidat nr./Candidate no. 10010

Dato/Date: 7/6.12 Side/Page: 27

Antall ark/Number of pages: 36

Emnekode/Subject SØK 3001

Denne kolonnen er forbeholdt sensor

This column is for external examiner

c) Vi vil teste $B_3 = B_1$ (lik effekt av pris og produktivitet. Og teste $B_3 = B_1$ vil være ekvivalent med å teste $B_1 - B_3 = 0$. Har da at vi ønsker å teste $H_0: B_1 - B_3 = 0$

$$\underbrace{B_1 - B_3}_{\theta} = 0$$

Vi vil da gjennomføre en transformasjon av den opprinelige likningen (som jeg antar nå er likning 3).

$$\ln W_t = \alpha \ln W_{t-1} + B_1 \ln p_t + B_3 \ln z_t + B_5 \ln U_{rt} + B_0$$

Det vi da gjør er å legge til og trekke fra $B_3 \ln p_t$, slik at vi får den transformerte likningen:

$$\ln W_t = \alpha \ln W_{t-1} + \underbrace{(B_1 - B_3)}_{\theta} \ln p_t + B_3 (\ln z_t + \ln p_t) - B_5 \ln U_{rt} + B_0$$

Ser at denne likningen vil være ekvivalent med likning 4, hvor vi har definert en ny paramet $\theta = B_1 - B_3$

Har nå den informasjonen vi trenger for å kunne teste hypotesen vei. Siden vi her har kun en restriksjon kan vi bruke en t-test, og følge tilsvarende fremgangsmåte som i b)

1) Hypotesen: $H_0: B_1 - B_3 = 0$ vs $H_a: B_1 - B_3 \neq 0$

2) $t_{obs} = \dfrac{-0{,}04}{0{,}08} = -0{,}5$

3) $t_{crit} = 1{,}98$ $\alpha = 0{,}05$ frihetsgrader $122 - 5 = 117$

4) Forkastningskriteriet: $|t_{obs}| < |t_{crit}|$

$|0{,}5| < |1{,}98|$

Den gule kopien beholder du/Please keep the yellow page

Tapir Uttrykk 05.2012

NTNU

Kandidat nr./*Candidate no.* 10010

Dato/*Date:* 7/6.12 Side/*Page:* 28

Emnekode/*Subject* SØK 3001 Antall ark/*Number of pages:* 36

Denne kolonnen er forbeholdt sensor

This column is for external examiner

ØL

Vi kan ikke forkaste H0. Dette tyder da på at pris og arbeidsledighet har lik effekt på nominell lønn.

Den gule kopien beholder du/*Please keep the yellow page*

Tapir Uttrykk 05.2012

NTNU

Kandidat nr./Candidate no. 10010

Dato/Date: 7/6.12 Side/Page: 29

Emnekode/Subject: Sosle 3001 Antall ark/Number of pages: 36

Denne kolonnen er forbeholdt sensor

This column is for external examiner

d) Fra likning 4 har vi at $\alpha = 0{,}47$. α representerer den estimerte verdien W_{t-1} (lagget nominell lønn, en periode tilbake). Vi kan tenke at α sier noe om graden av treghet i tilpasningen. Hvis α er høy tyder dette på at tidligere verdier av nominell lønn har stor innvirkning på tilpasningen, slik at vi får en "sluggish" tilpasning (altså vil ikke tilpasning til ny likevekt skje øyeblikkelig). Videre ser vi at t-verdien til α er høy. Når vi tester om α er signifikant forskjellig fra 0 finner vi at Ho kan avises med god margin ($t^{obs} = 6{,}71$ som er en høy verdi. (Gjennomfører ikke noen offesiell test pga tidsbesparing, men har beskrevet hvordan dette gjøres i tidligere oppgaver)

Vi har altså at ~~wen~~ α har en signifikant effekt på W_t i modellen. Som videre da indikerer at ved en endring av variablene i modellen vil vi få en "sluggish" tilpasning. Merk at vi har at fortiden kan påvirke framtiden men ikke vise versa. Modellen vi har vi 4) kan beskrives som en autogressiv modell (lagget endogen variabel). Når vi har en slik modell kan det vises at $\alpha = 1 - \lambda$. For λ kan beskrives som justeringshastigheten. Har at en høy λ indikerer en rask tilpassning, mens en lav λ indikerer treghet i tilpasningen. Ser da at dette vil bety at en høy α vil gi treghet i tilpasning. Dette er fordi W_t vil være avhengig av W_{t-1} (tidligere verdier vil være av påvirkning). Samtidig ser vi at en lav α vil gi en "rask tilpasning" (rask bevegelse til ny likevekt). Siden det kun bes om en kort drøfting av α velger jeg derfor å ikke utlede hvordan vi kommer fram til resultatet $\alpha = 1 - \lambda$.

Har da at hvis nominell lønn øker med 1% i periode t-1 vil dette gi en økning på 0,47 prosent ~~i~~ i nominell lønn i periode t. Altså har vi at fortiden påvirker fremtiden

Den gule kopien beholder du/Please keep the yellow page

Denne kolonnen er forbeholdt sensor
This column is for external examiner

Vi ser at dette er en log-log spesifikasjon, og kan dermed betrakte parameterne som elasisiteter:

$$\frac{1}{W_t}\frac{\partial W_t}{\partial W_{t-1}} = \alpha \frac{1}{W_{t-1}} \Rightarrow \frac{W_{t-1}}{W_t}\frac{\partial W_t}{\partial W_{t-1}} = \alpha$$

Skal nå finne den kortsiktige og langsiktige effekten av

i) Økning i arbeidsledighetsraten med 1%

ii) En økning i ledighetsraten fra 3% til 4%

For å finne de langsiktige effektene antar vi at vi har en langsiktig likevekt (steady-state) hvor $\bar{W}_* = W_{t-1} = W_t$
Dette gir da følgende likning

$$\ln \bar{W}_*(1-\alpha) = \theta p_t + B_3(\ln z_t + \ln p_t) + B_5 \ln Ur_t + B_0$$

V

$$(*)\ \ln \bar{W} = \frac{\theta}{1-\alpha} p_t + \frac{B_3}{1-\alpha}(\ln z_t + \ln p_t) + \frac{B_5}{1-\alpha} \ln Ur_t + \frac{B_0}{1-\alpha}$$

i) Arbeidsledighetsraten Ur_t øker med 1%
Kortsiktig effekt vil være gitt med parameteret foran Ur_t som er B_5. Har altså at hvis arbeidsledighetsraten øker med 1% så vil nominell reduseres med 0,05%. Merk at vi også her har et log-log nivå, og kan tolkes i form av elasisiteter.

V

Den langsiktige effekten kan vi finne ved å ta utgangspunkt i (*) ovenfor:

$$\Delta W_t = \frac{B_5}{1-\alpha} \Delta Ur_t$$

√ Dette gir: $\frac{-0,05}{1-0,47} = -0,094$

Det vil si at den langsiktige effekten av 1% økning i arbeidsledighetsraten ~~av~~ vil gi en reduksjon i nominell lønn på -0,094%

ii) Ser nå på en økning i ledighetsraten fra 3% til 4%
Det som nå er viktig å merke seg er log spesifikasjonen.
Vi kan finne endringen i w_t på følgende måte.
Kortsiktig effekt på lønnen vil fortsatt være gitt ved B_5.

$$\Delta W_t = -0,05(\log(4) - \log(3))$$

$$= -0,05 \cdot 0,288 = \underline{-0,0144}$$

Dette betyr at når ledighetsraten endrer seg fra 3 til 4% vil den

1,4% kortsiktige effekten gi en reduksjon i lønn på 0,01%
Forskjellen fra resultaten av den kortsiktige effekten i i) er at ~~med~~
~~log form~~ ~~når vi har forklaringsvariabelen i log har det betydning~~ om
vi nå ikke kun ser på 1% endring men på en 1% endring fra 3 til 4. Som da betyr at vi tar med i betraktning ~~den~~ at arbeidsledighetsraten har økt relativt i forhold til perioden før økningen skjedde.

Langtidseffekten kan vi finne basert på resultatet fra ⊛

$$\Delta W_t = \frac{B_5}{1-\alpha} \Delta U r$$

Husk at siden modellen er oppgitt i log form har vi at

$$\Delta W_t = \frac{-0,05}{1-0,47}(\log(4) - \log(3))$$

Denne kolonnen er forbeholdt sensor

This column is for external examiner

$$\rightarrow \frac{-0{,}05}{1-0{,}46} \cdot 0{,}288 = -0{,}027$$

2,7% Ser at ved den langsiktige effekten vil en økning i arbeids-ledighetsraten fra 3 til 4% gi en reduksjon i lønn på 0,027%

e) Det kan tyde på at lønn og pris er endogent bestemt innenfor en simultan modell. Dette betyr at bestemmelsen av lønn og pris avhenger av hverandre, og bestemmes simultant. Dette betyr da at vi vil få et endogenitets-problem ~~for~~ hvis vi bruker MKM, fordi her vil p bli betraktet som en eksogen variabel. Når p i virkeligheten er en endogen variabel vil dette da bety at MKM-estimering vil gi forventningsskjev estimator.

Når lønn og pris er endogent bestemt vil det bety at prisen vil påvirke lønnen og lønnen vil påvirke prisen. Vi kan tenke oss at vi har en 2 veis kausalitet:

$p \to W$ og $W \to p$

Som argumentert for ovenfor kan vi ikke lenger bruke MKM til estimering da vi vil få forventningsskjev estimator for p, som gjør at tolkningen av den partielle effekten vil gi en misvisende indikasjon.

Videre i denne oppgaven vil jeg anta at følgende ligninger gjelder for å kunne gi en beskrivelse av hvordan vi kan ta hensyn til dette problemet.

1) $W_t = \alpha_1 p_t + B_{10} + B_{11} Z_1 + U_1$ (Antar at Z_1 og

2) $p_t = \alpha_2 W_t + B_{20} + B_{22} Z_2 + U_2$ Z_2 er eksogene variabler)

Likningene ovenfor kan beskrives som strukturligninger i en simultan modell. Skal nå først vise at p_t og W_t vil være endogene. Dette kan vi gjøre ved å bruke Cramers regel å finne redusert form likningene.

$$W_t - \alpha_1 p_t = B_{10} + B_{11} Z_1 + U_1$$

$$-\alpha_2 W_t - p_t = B_{20} + B_{22} Z_2 + U_2$$

På matriseform kan dette skrives som

$$\begin{pmatrix} 1 & -\alpha_1 \\ -\alpha_2 & 1 \end{pmatrix} \begin{pmatrix} W_t \\ P_t \end{pmatrix} = \begin{pmatrix} B_{10} + B_{11} Z_1 + U_1 \\ B_{20} + B_{22} Z_2 + U_2 \end{pmatrix}$$

$$W_t = \frac{\begin{vmatrix} B_{10} + B_{11} Z_1 + U_1 & -\alpha_1 \\ B_{20} + B_{22} Z_2 + U_2 & 1 \end{vmatrix}}{\begin{vmatrix} 1 & -\alpha_1 \\ -\alpha_2 & 1 \end{vmatrix}}$$

$$W_t = \frac{B_{10} + B_{11} Z_1 + U_1 - \alpha_1 B_{20} - \alpha_1 B_{22} Z_2 - \alpha_1 U_2}{1 - \alpha_1 \alpha_2}$$

$$W_t = \underbrace{\frac{B_{10} - \alpha_1 B_{20}}{1 - \alpha_1 \alpha_2}}_{\Pi_{10}} + \underbrace{\frac{B_{11}}{1 - \alpha_1 \alpha_2}}_{\Pi_{11}} Z_1 - \underbrace{\frac{\alpha_1 B_{22}}{1 - \alpha_1 \alpha_2}}_{\Pi_{12}} Z_2 + \underbrace{\frac{U_1 - \alpha_1 U_2}{1 - \alpha_1 \alpha_2}}_{e_1}$$

$$W_t = \Pi_{10} + \Pi_{11} Z_1 + \Pi_{12} Z_2 + e_t$$

Som vil være likningen for lønnen på redusert form

Tilsvarende finner vi P_t:

$$P_t = \frac{\begin{vmatrix} 1 & B_{10} + B_{11} Z_1 + U_1 \\ -\alpha_2 & B_{20} + B_{22} Z_2 + U_2 \end{vmatrix}}{\begin{vmatrix} 1 & -\alpha_1 \\ -\alpha_2 & 1 \end{vmatrix}}$$

√

$$P_t = \frac{B_{20} + B_{22} Z_2 + U_2 - \alpha_2 B_{10} - \alpha_2 B_{11} Z_1 - \alpha_2 U_1}{1 - \alpha_1 \alpha_2}$$

NTNU

Kandidat nr./Candidate no. 10010

Dato/Date: 7/6.12 Side/Page: 35

Emnekode/Subject SØK 3001 Antall ark/Number of pages: 36

Denne kolonnen er forbeholdt sensor

This column is for external examiner

$$P_t = \underbrace{\frac{B_{20} - \alpha_2 B_{10}}{1-\alpha_1\alpha_2}}_{\pi_{20}} - \underbrace{\frac{\alpha_2 B_{11}}{1-\alpha_1\alpha_2}}_{\pi_{21}} Z_1 + \underbrace{\frac{B_{22}}{1-\alpha_1\alpha_2}}_{\pi_{22}} Z_2 + \underbrace{\frac{U_2 - \alpha_2 U_1}{1-\alpha_1\alpha_2}}_{e_2}$$

$$P_t = \pi_{20} + \pi_{21} Z_1 + \pi_{22} Z_2 + e_2$$

Som er redusert form likningen for pt

Fra de redusert formlikningene for W_t og P_t er vi at ~~...~~ pt vil være ~~...~~ korrelert med U_1 (dette ser vi på uttrykket for e_2). Vi ser også at W_t vil være korrelert med U_2 (ser fra e_1). Dette indikerer altså at vi får forventningsskjeve estimatorer ved bruk av MKM. Formelt kunne vi vist denne korrelasjon ved $\mathrm{Cov}(W_t, U_2)$ og $\mathrm{Cov}(p_t, U_1)$, men har ikke tid til å gjennomføre denne utledningen her, og er heller ikke av stor betydning for den videre diskusjonen i oppgaven.

For å ta hensyn til problemet med at W_t og P_t bestemmes simultant kan vi enten ~~...~~ bruke Z_2 som et instrument for p_t i likning en, og Z_1 som en instrument for W_t i likning to. For å kunne bruke disse som instrumenter er det viktig at disse er eksogene det vil si at

$\mathrm{Cov}(Z_1, U_i) = \mathrm{Cov}(Z_2, U_i) = 0$

Samtidig må vi ha at Z_2 er korrelert med P_t og Z_1 er korrelert med W_t, noe vi ser i fra redusert formlikningen for de 2 variablene at de mest sannsynlig vil være.

Et annet alternativ for å ta hensyn til dette endogenitetsproblemet er å estimere redusert form likningene

$\hat{W}_t = \hat{\pi}_{10} + \hat{\pi}_{11} Z_1 + \hat{\pi}_{12} Z_2 + \text{støy}$

$\hat{P}_t = \hat{\pi}_{20} + \hat{\pi}_{21} Z_1 + \hat{\pi}_{22} Z_2 + \text{støy}$

NTNU

Kandidat nr./Candidate no. 10010

Dato/Date: 7/6.12 Side/Page: 36

Emnekode/Subject SØK 3001 Antall ark/Number of pages: 36

Denne kolonnen er forbeholdt sensor

This column is for external examiner

Dette vil være første steg av 2 stegs MKM som er metoden vi kan bruke for å ta hensyn til endogenitetsproblemet.
Vi har da at vi kan ~~bruke~~ ~~at~~ erstatte p_t i likning 1) med $\hat{p}_t$ og ~~finn~~ så bruke MKM på denne likningen, som da vil gi forventningsrette estimatorer
Tilsvarende kan vi bruke $\hat{w}_t$ i stede for w_t i likning 2 for så å utføre MKM.

Har nå vist to måter som vi kan korrigere for endogenitets-problemet mellom w_t og p_t. Utfordringen ved bruk av instrumentvariabelmetoden vil være å finne gode instrument for p_t og w_t. Det er viktig at disse oppfyller betingelsen om at de er ukorrelerte med restleddet samt korrelerte med variablene de skal være instrument for.
Jeg valgte i denne delen av oppgaven å ta utgangspunkt i en enkel simultan modell mellom p_t og w_t, dette er for å gjøre intuisjonen best mulig.
I vår modell må vi altså ta hensyn til at w_t og p_t bestemmes simultant, og kan korrigere for dette ved 2 stegs MKM eller Instrumentvariabel. For at dette skal være mulig med likningen for w_t inneholdem eksogene variabler som ikke inngår i likningen for p_t og vise versa.