

## Sensorveiledning 3514 / 8614 Anvendt økonometri, H2017

**Generell opplysning: Ekstern sensor har fått tilsendt de artiklene oppgaven bygger på samt faglærers presentasjon av disse. Kandidatene har ikke tilgang til disse *under eksamen*, men hovedinnholdet forutsettes likevel velkjent.**

### Oppgave 1

a) Hovedsaken er her at vi har en «behandlingsgruppe» som påvirkes av politikk/reform/hendelse og en kontrollgruppe som ikke påvirkes. Videre har vi observasjoner for outcome-variabelen før og etter behandling. Den grunnleggende forutsetningen denne strategien bygger på er at *utviklingen* i outcomevariabelen ville vært den samme i behandlings- og kontrollgruppen i fravær av behandling (forutsetningen om parallelle trender). Videre er det sentral forutsetning at behandling er eksogen. Under disse forutsetningene kan vi finne kausaleffekten som endringen i forskjellen mellom behandlings- og kontrollgruppen der vi har beregnet gjennomsnittene for de to gruppene før og etter behandling. Rent teknisk kan vi alternativt finne kausaleffekten ved å estimere dummyregresjonen

$$(1) y = \beta_0 + \beta_1 DT + \beta_2 DA + \delta DA \cdot DT + u$$

der DT er dummy lik 1 for treatment og DA er dummy lik 1 for perioden etter treatment. Begrunn her at  $\delta$  tolkes som kausaleffekten.

Det er naturlig å peke på at ligning (1) kan utvides med kontrollvariable som (i) tar hensyn til eventuelt brudd på forutsetningen om parallelle trender, (ii) benyttes i robusthetssjekker og (iii) kan gi mer presise estimat.

DiD kan gjennomføres både ved bruk av «pooled cross sections» og paneldata i streng forstand. Paneldata gir imidlertid mer informasjon som f.eks. kan utnyttes ved å inkludere individspesifikke faste effekter. Treghet i tilpasningen kan analyseres hvis vi har info om flere år post treatment. Inkluder tidsdummy og interaksjonsledd for hvert enkelt år og undersøk hvordan parameteren foran interaksjonsleddet utvikler seg over tid.

b) Hypotesen som testes i artikkelen av Marianne Haraldsvik er om prestasjonsbasert opptak til videregående skole gir bedre elevprestasjoner i ungdomsskolen sammenlignet med bostedsbasert opptak. Hun utnytter her en reform i Hordaland med overgang fra bosteds- til prestasjonsbasert opptak. Elever i Hordaland utgjør behandlingsgruppen mens elever i andre fylker som opprettholder bostedsbasert opptak gjennom hele observasjonsperioden. Utgangspunktet vil her være en DiD-spesifikasjon som formulert i ligning (1) der y er elevprestasjoner i ungdomsskolen og  $DT = 1$  for elever i Hordaland. Argumenter for at «treatment» kan betraktes som eksogent siden det er fylkeskommunen som velger opptaksregime mens primærkommunen er ansvarlig for grunnskole.

Bør nevnes / diskuteres hvilke type kontrollvariable som i tillegg inkluderes og hvilke typer robusthetssjekker som foretas (*noe* av dette bør være med). Det siste gjelder (i) placebotesting / sjekker utvikling «pre treatment», (ii) separat estimering for «academic track», separat estimering for offentlige skoler, (iii) separate tidsdummyer og interaksjonsledd «pre treatment» for å teste om effekten øker over tid, (iv) sensitivitetssjekk ved å utelate enkeltkommuner.

b) Artikkel av Jon Fiva

$$(1) M_{it}^r = \delta_i^r + \theta_{it} + \beta b_{it} + \varepsilon_{it}^r$$

$$(2) M_{it}^n = \delta_i^n + \theta_{it} + \varepsilon_{it}^n$$

Kommenter at for hver gruppe (mottakere av sosialhjelp er treatment, ikke-mottakere er kontroll) inkluderes separate kommunespesifikke faste effekter, tidsvarierende kommunespesifikke effekter *som antas like mellom de to gruppene*. Sosialhjelpssatsen antas kun å påvirke netto-innflytting av mottakere.

Differensiering av (1) og (2) fjerner de to kommunespesifikke faste effektene (deltaledene), men  $\Delta\theta_{it}$  er fortsatt med i de differensierte motstykkene til (1) og (2). Ved å ta differensen mellom differensiert M for treatment og kontroll kommer Fiva fram til DiD-spesifikasjonen

$$(3) \Delta M_{it}^r - \Delta M_{it}^n = \beta \Delta b_{it} + u_{it}, \quad u_{it} = \Delta \varepsilon_{it}^r - \Delta \varepsilon_{it}^n$$

Gitt spesifikasjonene i (1) og (2) avhenger altså forskjellen i endring i de to migrasjonsratene kun av endringen i stønadssatsen (pluss restleddet). Forutsetningen om «parallele trender» som DiD-strategien bygger på kan knyttes til forutsetningen om at de tidsvarierende kommunespesifikke komponentene,  $\theta_{it}$ , er identiske mellom treatment- og kontrollgruppen.

Dersom utelatte, tidsvarierende, kommunespesifikke variable har ulik effekt på flyttebeslutningen for mottakere og ikke-mottakere betyr dette brudd på denne forutsetningen. I analysen foretar Fiva sensitivitetssjekker ved å inkludere et sett av observerbare variable. Videre kan IV-estimering også sies å ta hensyn til eventuelle endogenitetsproblemer ved utelatte kontroller.

d) Her er det naturlig å argumentere for at økt innflytting av mottakere gir kommunen et insentiv til å redusere satsen på sosialhjelp. Medfører toveis kausalitet (med motsatt fortegn) og at OLS anvendt på (3) gir bias mot null. (Kan gjøre dette formelt ved å formulere 2-ligningsmodell, men god diskusjon er like bra).

Fiva estimerer derfor (3) ved bruk av IV/2SLS. Ved instrumentering utnyttes en reform der «sentrale myndigheter» i februar 2001 innførte en norm for sosialhjelpssatsen. Fiva definerer 3 IV: Dummy lik 1 for kommuner der  $b <$  normen, nivået på  $b$  før reformen og interaksjonsledd mellom dummy og lagget  $b$ . Kjører 1.steg med  $b(2001)$  mot de tre instrumentene og estimerer i andresteget (3) for 2001 der faktisk  $b(2001)$  «erstattes» med predikert verdi.

Kandidatene bør videre diskutere gyldighet av IV (kan testes ved Sargan eller Hausman tester) og relevans (tester om IV har simultant signifikant effekt i førstesteget).

Kan også nevne / diskutere sensitivitetssjekker som foretas i artikkelen eller komme tilbake til dette under neste delspørsmål.

e) Main points: (i) OLS i (1) gir lavere estimert effekt av sosialhjelp enn estimering ved 2SLS (naturlig å sml med (3)). Rimelig siden det kan argumenteres for at simultanitet gir negativ bias ved OLS), (ii) Sammeligning av (2) og (3) indikerer at effekten er noe sterkere og mer signifikant for kortdistanseløpere (flytting mellom kommuner innen samme fylke). (iii) 2SLS er robuste overfor både inkludering av kontrollvariable (tar hensyn til at flytting, i første

rekke for kontrollgruppen, kan påvirkes av inkluderte kontroller) og valg av IV (kolonne 5). Kommenter at det benyttes sterke IV (F-verdier over +/- 50 i 2-4), og at bruk av kun en IV gir nesten likt estimat som ved bruk av alle tre IVs (kan tolkes som en practical Hausman-test).

## Oppgave 2

a) Hovedsak her: Interessevariabelen påvirkes diskontinuerlig av «assignment» / «forcing» variabel. Diskontinuerlige sprang i interessevariabelen gir eksogen variasjon rundt «cut-off» / terskelverdier som utnyttes når en ønsker å estimere kausaleffekten av denne på outcome. Poeng å undersøke særskilt om outcome påvirkes nær cutoff ved bruk av ulike vindusbredder. Argumenter for at «treatment» og «kontroll» sannsynligvis er mer like i andre henseender like over og like under cutoff. Videre poeng er at assignmentvariabelen i de fleste tilfeller må antas å påvirke outcome. Derfor viktig å kontrollere for denne, og å ta hensyn til potensielle ikke-lineariteter (ofte ved polynom) i kontrollfunksjonen. Generelt argument: Interessevariabelen er diskontinuerlig mens andre variable som oftest er kontinuerlige. Kan nevnes at en bør sjekke dette: Er det andre variable som også viser diskontinuitet i omegnen av terskelverdiene for interessevariabelen? I så fall alvorlig utfordring for RDD.

«Sharp design»: Interessevariabelen er en deterministisk funksjon (og bestemmes entydig) av assignmentvariabelen. «Fuzzy design»: Interessevariabelen betraktes som stokastisk, men ssh for diskontinuitet påvirkes av assignmentvariabelen nær potensielle cut-offs. Bør nevnes at ved fuzzy benyttes en 2steps metode der interessevariabelen instrumenteres.

Hensiktsmessig at kandidatene i denne delen trekker på en eller flere anvendelser på pensum. I gjennomgående eksempel i Lee og Lemieux, og i Petterson-Lidbom (2008) er interessevariabelen binær. I artiklene til Angrist og Lavy samt Leuven, Oosterbeek og Rønning (klassestørrelse) endres interessevariabelen diskontinuerlig, men ved flere terskelverdier på assignment, i Petterson-Lidbom (2012), neste spørsmål i oppgaven, endres antall representanter i kommunestyrene sprangvis som funksjon av folketall – flere cut-offs, deterministisk for Finland, ikke helt entydig for Sverige (fuzzy design).

Hvilke anvendelser som utnyttes er her en smaksak, men fordel å eksemplifisere forskjellen mellom sharp og fuzzy design med basis i en av disse. Kan også være hensiktsmessig å se 2a) og 2b) i sammenheng.

b) Med grunnlag i opplysningene gitt i oppgaveteksten bør kandidatene relativt raskt konkludere med at for Finland er antall representanter en deterministisk funksjon av innbyggertall så her har vi sharp design. For Sverige gjelder minimumskrav for antall representanter innenfor oppgitte intervall. Her har vi derfor fuzzy design.

Grunnmodellen som estimeres – både for Finland og Sverige er gitt ved:

$$(1) Y_{it} = \alpha_i + \lambda_t + \beta Csize_{it} + f(x_{it}) + u_{it}$$

der  $y$  er log til per capita kommunale utgifter og  $Csize$  log til antall representanter. Kommenter at det inkluderes kommunespesifikke faste effekter, faste tidseffekter samt ulike polynom på kontrollfunksjonen, det siste dels for robusthetssjekker.

For Finland estimeres (1) direkte ved bruk av OLS, dels basert på alle observasjoner, dels ved bruk av observasjoner innenfor ulike vindusbredder nær terskelverdiene.

For Sverige instrumenteres Csize ved å inkludere et sett av dummyvariable for intervallene som framgår av oppgaveteksten i førstestegsregresjonen. Andresteget estimerer (1) med «predikert Csize i stedet for faktisk Csize). Førstesteget inkluderer også kommunespesifikke faste effekter, tidseffekter og ulike polynom på kontrollfunksjonen.

Kan nevnes at en fordel med å studere kommuner innen et land er at samme institusjonelle forhold gjelder, relativt like kulturelle forhold mellom kommuner etc. Dette i motsetning til tverrsnittstudier basert på ulike land.

c) Fordelen ved inkludering av kommunespesifikke faste effekter at disse generelt kontrollerer for utelatte variable, ulike mellom kommuner, men faste over tid. Er et særlig poeng her når Petterson-Lidbom benytter alle observasjoner. Ulempen er generelt at ved en slik spesifisering må det være tilstrekkelig tidsvariasjon i interessevariabelen for at en skal kunne identifisere kausaleffekten av denne. I gjeldende setting betyr dette at det må være et tilstrekkelig antall kommuner der Csize over tid endres fra et intervall til annet.

d) Resultatene for Finland framstår som robuste både hva angår spesifisering av kontrollfunksjonen (saml (1) og (2)) og om hele sampelet benyttes eller om det kun benyttes observasjoner nær terskelverdiene (kolonne (3)). Overalt signifikant negativ effekt. For Sverige synes det å ha stor betydning for estimert effekt hvordan kontrollfunksjonen spesifiseres. Videre er estimerte effekter ikke-signifikante, og F-verdien fra førstesteget indikerer at instrumentene ikke er spesielt sterke (påpek gjerne sammenheng her mellom svake / moderate IV og uskarpt estimerte effekter i andreteget).

Ville konkludert med høy grad av intern validitet for Finland (robuste resultater) samt ekstern validitet i hvert fall når det gjelder retning på effekten siden forfatteren også finner (robust) negativ effekt for Sverige. Hva angår størrelsen på effekten kan en diskutere grad av ekstern validitet siden punkttestimatene for Sverige varierer mye mht spesifisering og heller ikke er statistisk signifikante.