

Sensorveiledning SØK3515 Vår 2021

a) Vi har her en standard problemstilling med binært valg / utfall. Kan starte med den lineære sannsynlighetsmodellen der relevante parametre estimeres ved OLS og parametrene tolkes direkte som marginale effekter på ssh for at individet er organisert. Argumenter for og formuler Logit og Probitmodellen. Forklar hvordan parametrene nå kan estimeres ved sannsynlighetsmaksimering. Minimum er her at kandidatene formulerer likelihoodfunksjonen og at maksimering av denne gir et sann av ikkelineære ligninger som løses for å finne parameterestimatene. Forklar hvordan marginaleffektene finnes når forklaringsvariabelen er kontinuerlig (alder) og når vi har dummy (kjønn). Forklar hvordan disse marginaleffektene avhenger av verdier på forklaringsvariablene og alle estimerte parametre i tillegg til estimert parameter foran hhv alder og kjønn.

Diskuter generelt fordeler og ulemper ved LMP versus de to ikkelineære alternativene.

b) Formuler modellen utvidet med interaksjonsledd $\text{kjønn} * \text{alder}$ og $\text{kjønn} * \text{utdanning}$. Formuler nullhypotesen formelt og forklar hvordan denne kan testes ved en likelihood ratio test. Gi gjerne intuisjon på testprosedyren.

c) Skill mellom tilfellet med og uten individuell heterogenitet. I siste tilfellet rett fram estimering av paneldatamodel analog til tverrsnittsmoell. I første tilfellet forventes forklaring av betinget logit. Pensum Verbeek 10.7.1 – 10.7.2. Få fram at prosedyren nå $T = 2$ strengt tatt gir en tverrsnittsmoell der forklaringsvariablene inngår på endringsform. Også viktig å få fram at kun observasjoner der individet endrer status gir info og utnyttes ved estimering. Observasjoner for individ som er organisert (uorganisert) i begge /alle perioder gir ingen nyttig info

d) Standard paneldatamestimering uten og med fixed effects. Få fram at fixed effectsmodellen kontrollerer for uobservert heterogenitet (konstant over tid), mens OLS på utransformerte variable gir bias pga manglende kontroll for slikt. Diskuter i forhold til oppgitte opplysninger om hvilke variable som inkluderes. Få fram at FE / WG kun utnytter variasjon over tid og at denne må være ulik for ulike individ. Standard OLS utnytter både tids- og tverrsnittsvariasjon. Diskuter i forhold til oppgitte opplysninger: Hvilke variable kan tenkes å variere mye, lite og ikke i det hele tatt (over tid). Hvilke konsekvenser har dette for valg av metode?

e) Tidsdummyer kontrollerer for utelatte aggregerte variable. Greit, men har dette betydning for estimering av effekter av interessevariablene? Ja, hvis inkluderte variable kan tenkes korrelert med utelatte makrovariable. Er dette sannsynlig i vårt tilfelle? Diskusjon av dette bør være med.

f) Her har vi et tilfelle med potensielt «endogen behandling» som er dekket av Verbeek side 269-274. Kan ta utgangspunkt i idealsituasjonen med kontrollert, men argumenter for at forutsetningen om «conditional independence» er svært restriktiv i forhold til vår problemstilling. Begrunn bias når denne ikke gjelder (endogen behandling). Forklar deretter Heckman-estimatoren for håndtering av problem med endogen behandling. Viktig å her få fram hva som bidrar til identifikasjon.

Siden dette er hjemmeeksamen med alle hjelpemidler tilgjengelig legges vekt på relevans og forståelse og at kandidatene svarer på spørsmålene i lys av de aktuelle problemstillinger.