

Eksamensopgave 3: Instrumentvariable

Betragt følgende model:

$$\log(\text{earnings}) = \beta_0 + \beta_1 s + \beta_2 \text{wexp} + \beta_3 \text{male} + \beta_4 \text{ethblack} + \beta_5 \text{ethhisp} + u \quad (1)$$

hvor *earnings* er timeløn i US dollars, *s* er uddannelse målt i antal års skolegang, *wexp* er erhvervserfaring målt i antal år, *male* er en dummy for køn, *ethblack* og *ethhisp* er racedummier for hhv. afroamerikanere og hispanics.

Vi har desuden tre instrumenter, moderens uddannelse målt i år (*sm*), faderens uddannelse målt i år (*sf*) og antal søskende (*siblings*).

Datasættet *data3*, som er tilgængelig på Moodle, indeholder disse variable målt for 520 amerikanere.

Nedenfor er der en række opgaver der skal løses. I forbindelse med de enkelte opgaver forventes det at der redegøres for den relevante teori. Det er altså ikke tilstrækkeligt blot at præsentere et "facit" for hver opgave.

Opgaver

1. Estimer modellen vha. OLS og kommenter på resultaterne.
2. Hvorfor kunne vi være bekymrede for at uddannelse er endogen?
3. Er *siblings*, *sm* og *sf* brugbare som instrumenter?
4. Test om uddannelse er endogen.
5. Estimer modellen vha. 2SLS hvor du gør brug af de tre beskrevne instrumenter. Sammenlign med resultaterne i spørgsmål 1.
6. Udfør overidentifikationstestet. Hvad konkluderer du?
7. Udfør hele analysen igen hvor du kun bruger *sm* og *sf* som instrumenter. Ændrer det på dine konklusioner?