

# Econometrie: tussentijdse test 2

Voor deze test werken we opnieuw met de Belgische budgetenquête uit 2000 (3816 huishoudens), beschikbaar in de folder

\\stud-shares.icts.kuleuven.be\data\GDhaene\Econometrie\Test2

Maak een folder Test2 aan op jouw bureaublad en kopieer budgetenquete.dta en test2.do daarin. De dataset bevat ondermeer de volgende variabelen:

*nadult*: aantal volwassenen

*nchi1*: aantal kinderen  $\leq 3$  jaar

*nchi2*: aantal kinderen  $> 3$  en  $\leq 6$  jaar

*nchi3*: aantal kinderen  $> 6$  en  $\leq 13$  jaar

*nchi4*: aantal kinderen  $> 13$  jaar

*eclo*: uitgaven voor kleding (in 2000, in euro)

*etot*: totale uitgaven voor niet-duurzame goederen (in 2000, in euro)

*distr*: gewest waarin het gezin woont

1: Brussel

2: Vlaanderen

3: Wallonië

*sexrp*: geslacht van het gezinshoofd (“rp” = reference person)

1: man

2: vrouw

*agerp*: leeftijd van het gezinshoofd

*edurp*: opleidingsniveau van het gezinshoofd

0: onbekend

1: geen diploma

2: diploma basisonderwijs

3: diploma secundair onderwijs (1ste cyclus)

4: diploma secundair onderwijs (2de cyclus)

- 5: diploma hoger onderwijs
- 6: diploma universitair onderwijs
- socrp*: sociale status van het gezinshoofd
- 0: zelfstandige
- 1: werkend
- 2: werkloos (nooit gewerkt)
- 3: werkloos (heeft vroeger gewerkt)
- 4: gepensioneerd
- 5: persoon met beperking
- 6: vervroegd pensioen
- 7: huisvrouw
- 8: baby, scholier, student
- 9: op andere wijze inactief

*sexpa, agepa, edupa, socpa*: analoog aan *sexrp, agerp*, etc., maar dan voor de partner van het gezinshoofd (indien aanwezig)

Beantwoord de volgende vragen. Schrijf de antwoorden neer op het antwoordblad. Je mag gebruik maken van test2.do en dat bestand verder aanvullen.

1. Beschouw de regressievergelijking

$$w = \beta_0 + \beta_1 \log(etot) + \beta_2 nadult + \beta_3 nchi1 + \beta_4 nchi2 + \beta_5 nchi3 + \beta_6 nchi4 + u$$

waarbij  $w$  het budgetaandeel is van kleding in de totale uitgaven aan niet-duurzame goederen:

$$w = \frac{eclo}{etot}.$$

- (a) Schat deze regressie met KK en rapporteer  $\hat{\beta}_1$  en zijn h-robuste standaardfout.

- (b) Veronderstel dat  $w = 0.05$  (d.w.z. 5% van de totale uitgaven gaan naar kleding). Veronderstel vervolgens dat  $etot$  met 10% stijgt en dat de gezinssamenstelling dezelfde blijft. Bereken het verwachte effect op  $w$  en op de uitgaven voor kleding,  $eclot$ . (Hint: dit is een vraag om “op papier” op te lossen, niet met een Stata commando.)
- (c) Het is niet moeilijk om na te rekenen dat de inkomenselasticiteit van de vraag naar kleding gegeven wordt door

$$\varepsilon = \frac{\partial eclot}{\partial etot} \times \frac{etot}{eclot} = 1 + \frac{\beta_1}{w}$$

(je hoeft dit niet na te rekenen). Dit impliceert dat

$$\begin{aligned}\varepsilon &< 1 \iff \beta_1 < 0, \\ \varepsilon &> 1 \iff \beta_1 > 0, \\ \varepsilon &< 0 \iff \beta_1 < -w.\end{aligned}$$

Is, volgens de geschatte  $\beta_1$ , kleding (i) een inferieur goed; (ii) een normaal en noodzakelijk goed; of (iii) een luxegoed? Motiveer kort.

- (d) Toets de hypothese dat de leeftijd van de kinderen geen effect heeft op  $w$ . (Hint: dit betekent *niet* dat het aantal kinderen geen effect heeft op  $w$ .)
- (e) Beschouw nu de regressie

$$\begin{aligned}w = & \beta_0 + \beta_1 \log(etot) + \beta_2 nadult \\ & + \beta_3 nchi1 + \beta_4 nchi2 + \beta_5 nchi3 + \beta_6 nchi4 + \beta_7 vrouw \\ & + \beta_8 edu2 + \beta_9 edu3 + \beta_{10} edu4 + \beta_{11} edu5 + \beta_{12} edu6 \\ & + \beta_{13} vla + \beta_{14} wal + u\end{aligned}\tag{1}$$

waarbij  $vrouw, edu2, edu3, \dots, edu6, vla, wal$  allemaal 0/1 variabelen zijn:

$$\begin{aligned}vrouw &= 1 \text{ als } sexrp = 2 \text{ (gezinshoofd is een vrouw)} \\ edu2 &= 1 \text{ als } edurp = 2 \text{ (gezinshoofd heeft opleidingsniveau 2)} \\ edu3 &= 1 \text{ als } edurp = 3 \text{ (gezinshoofd heeft opleidingsniveau 3)} \\ &\text{etc.} \\ vla &= 1 \text{ als } distr = 2 \text{ (gezin woont in Vlaanderen)} \\ wal &= 1 \text{ als } distr = 3 \text{ (gezin woont in Wallonië)}\end{aligned}$$

Schat (1) met KK, waarbij

- je de observaties waarvoor  $edurp = 0$  (d.w.z. opleidingsniveau is onbekend) uitsluit; doe dit met

... if `edurp` > 0 ...

(cf. lab sessie 1);

- je de 0/1 variabelen niet expliciet aanmaakt, maar in de regressie opneemt als `i.sexrp i.edurp i.distr` (cf. lab sessie 3).

Schrijf het Stata commando voor deze schatting voluit neer op je antwoordblad. Rapporteren van de schattingsresultaten hoeft niet.

- (f) Waarom voegen we in (1) geen 0/1 regressoren toe voor *man*, voor  $edurp = 1$  (geen diploma), en voor  $distr = 1$  (Brussel)?
- (g) Waarom zou het geen zin hebben om in (1) *vla* en *wal* te vervangen door één enkele regressor, namelijk *distr*?
- (h) Zou het zin hebben *edu2*, *edu3*, ..., *edu6* te vervangen door één enkele regressor, namelijk *edurp*? Zoja, is daar een bijkomende veronderstelling voor nodig (en zoja, welke)?
- (i) Beoordeel de volgende 3 uitspraken telkens met ja of nee, zonder uitleg. De schattingsresultaten tonen aan dat, ceteris paribus (d.w.z., terwijl de andere regressoren constant blijven)
  - vrouwen meer uitgeven aan kleding dan mannen;
  - hoger opgeleiden meer uitgeven aan kleding dan lager opgeleiden;
  - Vlamingen meer uitgeven aan kleding dan Brusselaars en Walen.
- (j) Voorspel hoeveel een alleenstaande Vlaamse vrouw zonder kinderen, met een universitair diploma en een jaarinkomen van 50000 euro, uitgeeft aan kledij (in euro).

2. We willen de beslissing modelleren van vrouwen om al dan niet te werken (in 2000), gegeven dat het gezin een man-vrouw koppel is (met of zonder kinderen). Beperk daartoe de steekproef door de volgende condities op te leggen:

- het gezin heeft 2 volwassenen; het gezinshoofd is een man, de partner is een vrouw (de dataset is anno 2000: mannen waren toen automatisch “gezinshoofd”);
- de vrouw is ouder dan 20 en jonger dan 60;
- de variabele *socpa* heeft de waarde 0, 1, 2, 3 of 7;
- de variabele *edupa* niet gelijk is aan 0.

Leg die condities op met

```
keep if (nadult == 2 & sexrp==1 & sexpa==2 & agepa>20 & agepa<60 & ...
      ... (socpa<=3 | socpa==7) & edupa!=0)
```

Creëer de variabele *werk* die de waarde 1 als de vrouw werkt of werkloos is ( $socpa \leq 3$ ) en 0 als de vrouw huisvrouw is ( $socpa = 7$ ). Beschouw nu het probit model

$$\begin{aligned} \Pr[\textit{werk} = 1|X] = \Phi(&\beta_0 + \beta_1 \textit{agepa} + \beta_2 \textit{agepa}^2 \\ &+ \beta_3 \textit{nchi1} + \beta_4 \textit{nchi2} + \beta_5 \textit{nchi3} + \beta_6 \textit{nchi4} \\ &+ \beta_7 \textit{edu2} + \beta_8 \textit{edu3} + \beta_9 \textit{edu4} + \beta_{10} \textit{edu5} + \beta_{11} \textit{edu6} \\ &+ \beta_{12} \textit{vla} + \beta_{13} \textit{wal}) \end{aligned}$$

waarbij  $X$  de gegeven lijst van regressoren is.

- Schat het model met de maximum likelihood schatter en rapporteer  $\hat{\beta}_{11}$  en zijn standaardfout.
- Leid een hoger diploma tot een grotere kans op werken? Motiveer kort.
- Bereken de voorspelde waarden van de kans op werken. Rapporteer daarvan het minimum, maximum en gemiddelde. Vergelijk ook met het gemiddelde van de variabele *werk*.
- Wat is het gemiddelde regionaal effect in Vlaanderen (op de kans op werken) t.o.v. Brussel? En in Wallonië t.o.v. Brussel? (Hint: gebruik `margins`.)