

Eksamenssæt 4

Opgave 1 - Opstil en lineær regressionsmodel for *participation* hvor du bruger de beskrevne forklarende variable.

(a) - Estimer modellen vha. OLS og kommenter på resultaterne.

MANGLER FORKLARING AF LPM + SVAGHEDER U ER PER DEFINITION HETEROSKEDASTIC
- DER BRUGES ROBUST SE

```
model_ols = lm(participation ~ income + age + agesq + educ + youngkids + oldkids + foreign, data = data)
robust_ols = coeftest(model_ols, vcov = vcovHC(model_ols, type = "HCO"))

screenreg(list(OLS = model_ols, OLS_robust_se = robust_ols), digits = 4)
```

```
##
## =====
##              OLS              OLS_robust_se
## -----
## (Intercept)  -0.3686         -0.3686
##              (0.2530)        (0.2358)
## income       -0.0035 ***     -0.0035 ***
##              (0.0007)        (0.0006)
## age          0.0634 ***       0.0634 ***
##              (0.0129)        (0.0119)
## agesq        -0.0009 ***     -0.0009 ***
##              (0.0002)        (0.0001)
## educ         0.0068           0.0068
##              (0.0060)        (0.0059)
## youngkids    -0.2390 ***     -0.2390 ***
##              (0.0314)        (0.0302)
## oldkids      -0.0475 **       -0.0475 **
##              (0.0172)        (0.0175)
## foreign      0.2572 ***       0.2572 ***
##              (0.0401)        (0.0401)
## -----
## R^2          0.1901
## Adj. R^2     0.1836
## Num. obs.    872
## =====
## *** p < 0.001; ** p < 0.01; * p < 0.05
```

```
#summary(model_ols)
```

Alle signifikante på 0,1% på nær oldkids på 1% og educ er ikke signifikant. INGEN FORSKEL PGA ROBUST SE

(b) - Test om den partielle effekt af uddannelse er forskellig fra nul.

For at teste hvorvidt den partielle effekt af en variabel er forskellig fra nul bruges en t-test. Hvorvidt nulhypotesen afvises afhænger af den beregnede t-score og dertilhørende p-værdi

$$H_0 : \beta_4 = 0$$

$$H_1 : \beta_4 \neq 0$$

T-scoren beregnes ud fra den estimerede β samt den tilhørende standardafvigelse. Dette kan gøres, da nulhypotesen er, at den faktiske værdi er nul, hvorfor dette led ikke indgår i formelen.

$$t = \frac{\hat{\beta}_j}{se(\hat{\beta}_j)}$$

I nedenstående vil de robuste standardafvigelser blive benyttet til udregningen af t-scoren

```
## Kritisk værdi ved 5% = 1.9626913
```

```
## Kritisk værdi ved 1% = 2.5814857
```

```
t = 0.0068/0.0059
```

```
## t-score = 1.1525424
```

(c) - Test om den partielle effekt af alder er forskellig fra nul.

T-test?

Opgave 2 - Opstil både en logit- og en probit-model for par t i cipat ion hvor du bruger de beskrevne forklarende variable.

(a) - Estimer modellerne.

```
screenreg(list("LPM OLS" = model_ols, Logit = logit, Probit = probit))
```

```
##
## =====
##               LPM OLS      Logit      Probit
## -----
## (Intercept)    -0.37      -4.39 ***   -2.67 ***
##               (0.25)      (1.30)      (0.78)
## income         -0.00 ***   -0.02 ***   -0.01 ***
##               (0.00)      (0.00)      (0.00)
## age            0.06 ***     0.33 ***     0.20 ***
##               (0.01)      (0.07)      (0.04)
## agesq          -0.00 ***   -0.00 ***   -0.00 ***
##               (0.00)      (0.00)      (0.00)
## educ           0.01         0.04         0.02
##               (0.01)      (0.03)      (0.02)
## youngkids      -0.24 ***   -1.18 ***   -0.71 ***
##               (0.03)      (0.17)      (0.10)
## oldkids        -0.05 **    -0.24 **    -0.14 **
##               (0.02)      (0.08)      (0.05)
## foreign         0.26 ***     1.19 ***     0.73 ***
##               (0.04)      (0.20)      (0.12)
## -----
## R^2            0.19
## Adj. R^2       0.18
## Num. obs.      872         872         872
## AIC            1032.15      1031.65
## BIC            1070.32      1069.82
## Log Likelihood -508.08      -507.83
## Deviance       1016.15      1015.65
## =====
## *** p < 0.001; ** p < 0.01; * p < 0.05
```

ESTIMATER FOR LOGIT OG PROBIT KAN IKKE FORTOLKES SOM DE ER

(b) - Test om den partielle effekt af uddannelse er forskellig fra nul.

(c) - Test om den partielle effekt af alder er forskellig fra nul vha. et likelihoodratio-test.

Opgave 3 - Vi vil gerne sammenligne den partielle effekt af *income* på tværs af modellerne. Beregn average partial effect (APE) og kommenter på resultaterne.

BRUGER ROBUST SE

```
ape_logit = logitmfx(logit, data = data, atmean=F, robust = T)
```

```
screenreg(list(ape_logit = ape_logit), digits = 4)
```

```
##
## =====
##                ape_logit
## -----
## income          -0.0046 ***
##                  (0.0010)
## age              0.0657 ***
##                  (0.0139)
## agesq            -0.0009 ***
##                  (0.0002)
## educ             0.0077
##                  (0.0060)
## youngkids        -0.2350 ***
##                  (0.0403)
## oldkids          -0.0470 **
##                  (0.0176)
## foreign          0.2466 ***
##                  (0.0409)
## -----
## Num. obs.        872
## Log Likelihood   -508.0766
## Deviance         1016.1533
## AIC              1032.1533
## BIC              1070.3196
## =====
## *** p < 0.001; ** p < 0.01; * p < 0.05
```

```
ape_logit
```

```
## Call:
## logitmfx(formula = logit, data = data, atmean = F, robust = T)
##
## Marginal Effects:
##              dF/dx      Std. Err.      z      P>|z|
## income    -0.004610926  0.001012942 -4.55201 0.0000053135052 ***
## age        0.065744335  0.013887098  4.73420 0.0000021991793 ***
## agesq     -0.000929881  0.000175837 -5.28832 0.0000001234475 ***
## educ       0.007705869  0.006036510  1.27654    0.2017634
## youngkids -0.235006974  0.040327957 -5.82740 0.0000000056299 ***
```

```
## oldkids    -0.046973245  0.017550880 -2.67640      0.0074417 **
## foreign     0.246583549  0.040923752  6.02544 0.0000000016865 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## dF/dx is for discrete change for the following variables:
##
## [1] "foreign"
```

Opgave 4 - Vi vil gerne sammenligne den partielle effekt af *foreign* på tværs af modellerne. Beregn APE og kommenter på resultaterne.

Opgave 5 - Hvorfor er APE at foretrække frem for partial effect at the average (PEA)?

Opgave 6 - Sammenlign modellernes evne til at prædiktere ved at beregne percent correctly predicted for hver model.