# WEEKEND PROBLEM

August 6, 2023

Hannah Massenbauer

202175166

Advanced Econometrics

# Q1: Kids

## a)

Drawbacks:

1. Observed outcome variable is between 0 and 1. That is neglected by OLS .
2. Heteroskedasticity issue -> bad for A4 -> loss of efficiency
3. error term is non-normal -> bad for A5 -> difficulties to draw inference
4. possible non-linear relations get ignored

Benefits:

1. easy to interpret and compute

Why its a linear probability model:
Because the model uses a binary outcome variable it will take the value one or not. It is still a usual OLS estimation, however due to the nature of the outcome variable has the above mentioned drawbacks. The model predicts the following:

$$E(y_i|x_i) = 1 * Pr(y_i = 1|x_i) + 0 * Pr(y_i = 0|x_i)$$

$$E(y_i|x_i) = Pr(y_i = 1|x_i)$$

Which resembles a linear probability of displaying $y_i = 1$ (in our case having kids). We observe in $\beta$ the change in the probability of having kids or not under ceteris paribus.

## b)

$H_0$ = test no diminishing effect of age = no effect

$H_A$ = there is a negative (diminishing) effect

$$H_0 : \beta_3 = 0$$

$$H_A : \beta_3 < 0$$

Since normality is violated I have to rely on asymptotic tests, where I would rely on a one sided z-test statistic and for the standard errors I would use robust ones to obtain a reliable test result (OLS is still inefficient!).

$$z = \frac{\hat{\beta} - \beta}{se(\hat{\beta}} \sim^{a} N(0,1)$$

Which only holds if the sample size is large.
The rejection rule: $|z| > z_\alpha$ This implies that if $\beta_3$ is far away from the null, we can reject the null.

The standard errors can be obtained by:
$SE_i = \sqrt{VA\hat{R}(\hat{\beta_i})}$
$VAR(\hat{\beta}_i)$ can be obtained by the $a_{ii}$ element of the variance covariance matrix, which would be the $i^{th}$ element of the diagonal of that matrix. The diagonal displays the variances and to obtain the standard error taking the root will yield the demanded standard errors.

## c)

The regression based regression is splitted in two steps:

1. Run the entire OLS regression and store the obtained residuals
2. Premultiply the dependent and the independent variable of interest only with the residual maker matrix obtained by the residuals from the first stage. As we have absorbed

everything what is explained by the other regressors we can achieve in the second stage the effect of education on kids, without the others influence.

## Q2: MLE and classical testing

### a)

Random sample how obtain $\hat{\theta}$:

Set a joint distribution, obtain log likelihood, FOC, set = 0, make $\theta$ explicit From the probability form of a density going to the statistic form of a Likelihood function I sum the terms over n and based on its Bernoulli form I rewrite it s.t. it will be later on easier to derive:

$$L = \sum_{i=1}^{n} (\theta) + \sum_{i=1}^{n} (1-\theta)^{y_i}$$

$$logL(\theta) = n * log(\theta) + \sum_{i=1}^{n} y_i * log(1-\theta)$$

From this I can easily set up the FOC:

$\hat{\theta}_{MLE} = argmaxlogL(\theta) = 0$

or

$$\frac{\mathrm{d}logL(\theta)}{\mathrm{d}\theta}\Big|_{\hat{\theta}_{\hat{\theta}MLE}} = 0$$

The MLE estimator aims to maximize the likelihood to observe the drawn data based on our imposed joint distribution. As a maximum is an extremum, we take the FOC and by the SOC < 0 we could show that its acutally a maximum. It is an iterative procedure until the maximum is approached and does not necessarily yield unbiased estimates, but very efficent ones (Cramer-Rao lower bound).

## b)

MLE is consistent but not unbiased as:

$$plim\hat{\theta}_{MLE} = \frac{1}{1+\bar{y}}$$

where

$$E(y_i) = \bar{y}$$

By using law of large numbers and Slutsky where we converge toward n->∞ the average of sample Y becomes $\bar{y}$ which is equal to the expected value of $y_i$. Therefore we can write:

$$plim\hat{\theta}_{MLE} = \frac{1}{1+\bar{y}} = \frac{1}{1+E(y_i)} = \frac{1}{1+\frac{1-\theta}{\theta}}$$

After bringing the denominator on the same line, I have 1 over $\theta$ left in the denominator and by multiplying the outer parts divided by the inner parts of the expression I obtain $\theta$. This implies that my MLE estimator converges in probability to the true parameter value.

For unbiasedness

$$E(\hat{\theta}_{MLE}) = \frac{1}{1+E(\tilde{y})}$$

would need to hold what is not the case by Jenses' Inequality. Therefore, the MLE estimator is consistent, but in general not unbiased.

## c)

Use unrestricted model and show that the coefficent for age is zero, therefore from $\hat{\theta} - \tilde{\theta}$ only the $\hat{\theta}$ is left in the test statistic. I use the coefficient to conduct a test:

$z = \frac{\beta_{age}}{se(\beta_{age})} \sim^a N(0,1)$ under $H_0$

the normal distribution is imposed as I accounted for heteroskedasticity in the error term by dividing through the standard errors.

After obtaining the second derivatives of the log L function I can construct a Hessian matrix. On the diagonal of the Hessian the variances are given, and by taking the square root of the variance I obtain the standard errors.

Having the coefficient and the standard errors allows me computing the test statistic, from which I can get the p-values which are a probability of how likely it is that my test statistic is larger than the critical value.

## d)

Likelihood ratio principle:

This test indicates the loss of fit of switching from the unrestricted to the restricted.

$$H_0 : \beta_{public} = \beta_{addon} = 0$$

$$H_A : \beta_{public} \neq 0, and/or, \beta_{addon} \neq 0$$

One has to compute the restricted and the unrestricted model and compare them afterward in the following test statistic:

$$LR = 2(lnL^R - lnL^u) -^d > \chi^2_2$$

if $LR > \chi^2_2$ then we can reject the null at the set significance level as the test statistic is larger than the critical value.

## e)

2: Interpretability of $\beta_k$

Taking the derivative of $\lambda_i$ wrt $x_k$ yields a marginal effect. This necessary step can be intuitively explained by the circumstance that the marginal effect is not constant but changing based on its level over x. If one wants the marginal effect at the means of the covariates one would need to multiply $\beta_i$ with the $exp(\bar{x}'\hat{\beta})$. But there are more concepts on how to calculate the marginal effect, and depending on the question it makes sense to consider other forms, such as the overall mean of marginal effects for example.