Master 1 Economie et Droit
Session 2: Econometrics (1 hour, no document authorized)

Exercise 1 (5 points)

We consider the model : (M1) $y_i = \beta_0 + \beta_1 z_i + \beta_2 w_i + \beta_3 x_i + u_i$, i=1, …, N, with *u* iid N(0, σ^2).

1. We want to test H0 :$\beta_1 - 2\beta_2 = 1$ vs H1 : $\beta_1 - 2\beta_2 \neq 1$. Give two ways to solve this problem.

We suspect now that $V(u_i) = \sigma^2 \exp(\alpha x_i)$ in (M1).

2. What is the problem? What are the consequences if we apply standard OLS?
3. Give the procedure to estimate correctly this model.

Exercise 2 (8 points)

We consider the model: $Y_i = \beta_0 + \beta_1 W_i + \alpha X_i + u_i \ ; i = 1, \ldots, N$, where *W* is an endogenous regressor, *X* is an exogenous regressors, and we consider 2 instrumental variables (IV) denoted Z_1 and Z_2.

1. What Gauss-Markov assumption does not hold with endogenous regressors?
2. Give 2 possible sources of endogeneity (explain).
3. What are the conditions for IV to be valid?
4. Describe a procedure to test overidentifying restrictions.
5. Give the detailed procedure to correctly estimate the model.

Exercise 3 (7 points)

We want to analyze the determinants for labor force participation of married women aged 35 years or less. We estimate a **Probit model** using a sample of 2172 women, where the dependent variable "*Y*" is a binary variable equal to 1 if the woman is in labor force, 0 otherwise, and the explanatory variables are: *educ* (years of schooling), *age*, *h_age* (husband age) and *kid* (=1 if at least one kid, 0 otherwise). The results are:

Dependent Variable: Y

Variable	Coefficient	Std. Error	z-Statistic	Prob.
AGE	0.039	0.008	4.875	0.0000
EDUC	0.106	0.012	8.833	0.0000
KID	-0.596	0.073	-8.164	0.0000
H_AGE	-0.027	0.005	-5.400	0.0000
C	-0.907	0.242	-3.748	0.0002

McFadden R-squared	0.071346	Mean dependent var	0.610958
S.D. dependent var	0.487645	S.E. of regression	0.465740
Akaike info criterion	1.245876	Sum squared resid	470.0515
Schwarz criterion	1.258959	Log likelihood	-1348.021
Hannan-Quinn criter.	1.250660	Deviance	2696.042
Restr. deviance	2903.172	Restr. log likelihood	-1451.586
LR statistic	207.1305	Avg. log likelihood	-0.620636
Prob(LR statistic)	0.000000		

Obs with Dep=0	845	Total obs	2172
Obs with Dep=1	1327		

1. Explain why a Probit model is preferred to a standard regression model?
2. In the case of one regressor, denoted *X*:
   a. write the Probit model and define the underlying latent model
   b. describe the estimation procedure
3. Using the table of results:
   a. comment these results
   b. what is the effect of *age* on the probability of participation to the labor force for a woman with 2 kids, with 14 years of schooling, and husband 35 years old?

(*Reminder*: the probability density function for X~N(0,1) is $f(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})$).