GENEVA GRADUATE INSTITUTE

INSTITUT DE HAUTES ÉTUDES INTERNATIONALES ET DU DÉVELOPPEMENT

GRADUATE INSTITUTE OF INTERNATIONAL AND DEVELOPMENT STUDIES

ECONOMETRICS I - MIE

Examples of Questions for the Final Exam[1]

Prof. Julia Cajal-Grossi - TA Viktoria Vidahazy

1. Consider the *two-stage least squares* estimator $\hat{\boldsymbol{\beta}}_{2SLS}$ as studied in class, using an $N \times L$ matrix $\boldsymbol{Z}$ of instruments (as usual $N > L$).

   (a) Assume $\boldsymbol{H}$ is an invertible $L \times L$ matrix of constants (fixed and finite), and define a new matrix of instruments, $\boldsymbol{Z}^* = \boldsymbol{ZH}$. Call $\hat{\boldsymbol{\beta}}_{2SLS^*}$ the *two-stage least squares* estimator using instruments $\boldsymbol{Z}^*$.

   - Construct this estimator.
   - Is it different from $\hat{\boldsymbol{\beta}}_{2SLS}$? Please justify your answer with an analytical derivation.

   (b) Is $\hat{\boldsymbol{\beta}}_{2SLS^*}$ a consistent estimator of $\boldsymbol{\beta}$? You may argue for this answer, without any formal proof. If you choose not to present a formal proof, please make sure you explain your reasoning in full.

2. You are given access to cross-sectional data on 3,005 individuals. For each individual, you observe the natural log of hourly wages (lhw), age measured in years (age), a dummy variable (female), which takes the value one if the individual is a female and is zero otherwise, the product of female and age (femage), years of work in the current job (tenure) and its square (tenure2). Two regressions are estimated, and produce the following output:

[1] This document contains examples that illustrate the *types* of questions that may appear in the final exam. This material is not exhaustive and should be used as general guidance only.

```
(1) reg lhw age female femage tenure tenure2

Source       |         SS   df        MS                   Number of obs = 3005
-------------+-----------------------------                F( 5, 2299) = 60.48
Model        | 100.000000    5     16.808223                  Prob > F = 0.0000
Residual     | 600.000000 2999    .277909507                 R-squared = 0.1429
-------------+-----------------------------                Adj R-squared = 0.1443
Total | 700.000000        3004   .313782583                  Root MSE = .52717
-------------------------------------------------------------------------------
lhw          | Coef.      Std. Err.    t       P>|t|     [95% Conf. Interval]
-------------+-----------------------------------------------------------------
age          | .0042398   .0014985    2.83     0.005     .0013012   .0071784
female       | -.0200000  .0500000   -0.40     0.745    -.1860868   .1331442
femage       | -.0100000  .0020022   -5.000    0.000    -.0098752  -.0020224
tenure       | .0060000   .0006000   10.00     0.000     .0021002   .0033886
tenure2      | -.0005000  .0001000   -5.00     0.000    -.0007000  -.0003000
_cons        | 1.764674   .0596079   29.60     0.000    1.647783   1.881564
-------------------------------------------------------------------------------


(2) reg lhw age female femage

Source       |     SS        df        MS                  Number of obs = 3005
-------------+-----------------------------                F( 3, 2301) = 65.24
Model        | 50.0000000    3    18.8903724                  Prob > F = 0.0000
Residual     | 650.000000 3001    .289562779                 R-squared = 0.0714
-------------+-----------------------------                Adj R-squared = 0.0772
Total        | 700.000000 3004 .313782583                      Root MSE = .53811
-------------------------------------------------------------------------------
lhw          | Coef.     Std. Err.     t         P>|t|    [95% Conf. Interval]
-------------+-----------------------------------------------------------------
age          | .0084819  .0014144     6.00       0.000     .0057082    .0112556
female       | -.049775  .082544     -0.60       0.547    -.2116433    .1120934
femage       | -.0056949 .0020203    -2.82       0.005    -.0096568    -.0017331
_cons        | 1.762608  .0586584    30.05       0.000    1.647579     1.877637
-------------------------------------------------------------------------------
```

(a) Assume the goal of the analysis is to understand the evolution of individuals' wages over time, and whether they follow a different path for females relative to male workers. Which of the two specifications above do you believe is best? Explain your reasoning.

(b) Do wages increase, decrease or remain unchanged with every year of experience in the job? If they change, by how much? Please include a quantitative interpretation of the results, establishing the relationship between wages and tenure.

(c) Test the hypothesis that the coefficients on tenure and tenure2 are jointly significant in the model.

(d) One may consider removing the variable female from the regressions. What do you expect the effect of this change to be, on the OLS estimates?

3. The following model characterizes the smoking behavior of individuals, in relation to their income:

$$Numcigs_i = \beta_0 + \beta_1 Income_i + u_i. \tag{1}$$

(a) You suspect the presence of measurement error in the left hand side (dependent) variable, which is the number of cigarettes smoked. In other words:

$$Numcigs^{\text{observed}} = Numcigs^{\text{true}} + e$$

where e is uncorrelated, simple measurement error.

Given the following information, please outline the consequences of this type of measurement error for the OLS estimation of (1). Explain your reasoning.

N = 100	$\text{Cov}\left(Numcigs, Income^{\text{true}}\right) = 5$	
	$\text{Cov}\left(Numcigs, Income^{\text{observed}}\right) = 1$	
	$\text{Var}\left(Income^{\text{true}}\right) = 5$	
	$\text{Var}(Numcigs) = 200$	
	$\text{Var}(e) = 100$	$\text{Var}(u) = 100$
	$\text{E}(u) = 0$	$\text{Cov}(e, u) = 0$
	$\text{E}(e) = 0$	

bias → attunnination bias

(b) Now consider that, instead, the right hand side variable is measured with error, as follows:

$$Income^{\text{observed}} = Income^{\text{true}} + w$$

where, again, w is simple measurement error. For this case, please find:

- the true (unobserved) β_1 capturing the relationship of income and cigarette consumption in the absence of measurement error.
- the OLS estimate that you would obtain with the data available (i.e. with the mismeasured regressor).

(c) Compare the two expressions and explain any differences you may see.

(d) You are asked to propose a solution for the measurement error problem. Please explain in detail.

4. True/False/Explain. To test for heteroskedastic errors in a linear model, it is useful to regress functions of the absolute values of least-squares residuals (e.g. the squared residuals) on functions of the regressors. The R-squared from this second stage regression will be (approximately) distributed as chi-square random variable under the null hypothesis of no heteroskedasticity, with degrees of freedom equal to the number of no-constant functions of the regressors in the second-stage.

5. Consider the following statement. *In applied work, it is common that the assumption of mean independence of the disturbance with respect to the matrix of regressors in the model is violated. When this is the case, OLS will not perform well as an estimator of the model parameters, but other estimators may be available.* Discuss this statement. In particular, explain what we mean by *"OLS will not perform well"*, and describe the alternative estimators that may be available. State and explain the conditions under which an alternative estimator is preferred.

F-test > 10 ⇒ Good instrument.

1. Consider the *two-stage least squares* estimator $\hat{\boldsymbol{\beta}}_{2SLS}$ as studied in class, using an $N \times L$ matrix $\boldsymbol{Z}$ of instruments (as usual $N > L$).

   (a) Assume $\boldsymbol{H}$ is an invertible $L \times L$ matrix of constants (fixed and finite), and define a new matrix of instruments, $\boldsymbol{Z}^* = \boldsymbol{Z}\boldsymbol{H}$. Call $\hat{\boldsymbol{\beta}}_{2SLS^*}$ the *two-stage least squares* estimator using instruments $\boldsymbol{Z}^*$.

   - Construct this estimator.
   - Is it different from $\hat{\boldsymbol{\beta}}_{2SLS}$? Please justify your answer with an analytical derivation.

   (b) Is $\hat{\boldsymbol{\beta}}_{2SLS^*}$ a consistent estimator of $\boldsymbol{\beta}$? You may argue for this answer, without any formal proof. If you choose not to present a formal proof, please make sure you explain your reasoning in full.

(1a) $\hat{\beta}_{2SLS} = (X' P_Z X)^{-1} (P_Z X)' y \qquad Z^* = ZH \Rightarrow$ Construct $\hat{\beta}_{2SLS}$ from Z^*

$P_Z^* = Z^* (Z^{*\prime} Z^*)^{-1} Z^{*\prime} = ZH(H'Z'ZH)^{-1} H'Z'$

$= Z\cancel{H H^{-1}} (Z'Z)^{-1} \cancel{(H')^{-1} H'} Z' = Z(Z'Z)^{-1} Z' = P_Z$

because invertable & H is invertable

$\Rightarrow \hat{\beta}_{2SLS^*} = (X' P_Z^* X)^{-1} (P_Z^* X)' y$

$= (X' P_Z X)^{-1} (P_Z X)' y$

$= \hat{\beta}_{2SLS}$

(1b) if $\hat{\beta}_{2SLS}$ is consistent $\Rightarrow \hat{\beta}_{2SLS^*}$ is consistent as well.

☆ Write down all the assumptions.

Derivation: $\hat{\beta}_{2SLS} - \beta = (X' P_Z X)^{-1} (P_Z X)' (X\beta + u) - \beta$

$= \cancel{(X' P_Z X)^{-1} (X' P_Z X)} \beta + (X' P_Z X)^{-1} X' P_Z u - \beta$

$= (X' P_Z X)^{-1} X' P_Z u$

$= \underbrace{[X'Z(Z'Z)^{-1} Z'X]^{-1}}_{①} \underbrace{[X'Z(Z'Z)^{-1} Z'u]}_{②}$

To show: $\text{plim}(\hat{\beta}_{2SLS} - \beta) = 0$

By D5: $\text{plim}\left(\frac{Z'Z}{n}\right) = Q_{ZZ}$ PD finite. Then Q_{ZZ}^{-1} will be the same

D6: $\text{plim}\left(\frac{Z'X}{n}\right) = Q_{ZX}$ finite matrix with rank of K ($L \times K$ & $K < L$)

$\nearrow$ in plim form

For ①: $A'BA$ with a B which is a PD & A is full column rank, $\Rightarrow A'BA$ is also a PD finite matrix.

For ②: $X'Z(Z'Z)^{-1}Z'u \Rightarrow$ We need $Z'u$: $\frac{1}{n}\sum z_i u_i = \bar{w}$

With C3: $E[u_i|z_i]=0$ Exogeneity $\Rightarrow E(\bar{w})=0 \Rightarrow plim\left(\frac{Z'u}{n}\right)=0$

$V(\bar{w}) = V[\underbrace{E(\bar{w}|Z)}] + E[\underbrace{V(\bar{w}|Z)}]$

$= 0 + E[V(\frac{1}{n}\sum z_i u_i | Z)]$

$= E[\frac{\sigma^2}{n} \frac{Z'Z}{n}]$

$\lim_{n\to\infty} \frac{\sigma^2}{n} = 0$ & $plim\left(\frac{Z'Z}{n}\right) = Q_{ZZ}$ P.D by D5.

$\Rightarrow Plim\, E[\frac{\sigma^2}{n} \frac{ZZ}{n}] = 0$

$\Rightarrow$ ① $\times$ ② $=1 \Rightarrow \hat{\beta}_{2SLS}$ is consistent

2. You are given access to cross-sectional data on 3,005 individuals. For each individual, you observe the natural log of hourly wages (lhw), age measured in years (age), a dummy variable (female), which takes the value one if the individual is a female and is zero otherwise, the product of female and age (femage), years of work in the current job (tenure) and its square (tenure2). Two regressions are estimated, and produce the following output:

→ 2999

```
(1) reg lhw age female femage tenure tenure2

Source       |          SS   df        MS                  Number of obs = 3005
-------------+-----------------------------                F( 5, 2299) = 60.48
Model        | 100.000000    5    16.808223                Prob > F = 0.0000
Residual     | 600.000000 2999    .277909507               R-squared = 0.1429
-------------+-----------------------------                Adj R-squared = 0.1443
Total | 700.000000        3004   .313782583                Root MSE = .52717
------------------------------------------------------------------------------
lhw          | Coef.       Std. Err.     t        P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
age          | .0042398    .0014985     2.83     0.005     .0013012    .0071784
female       | -.0200000   .0500000    -0.40     0.745    -.1860868    .1331442
femage       | -.0100000   .0020022    -5.000    0.000    -.0098752   -.0020224
tenure       | .0060000    .0006000    10.00     0.000     .0021002    .0033886
tenure2      | -.0005000   .0001000    -5.00     0.000    -.0007000   -.0003000
_cons        | 1.764674    .0596079    29.60     0.000    1.647783    1.881564
------------------------------------------------------------------------------
```

Diminishing marginal scale

```
(2) reg lhw age female femage

Source       |     SS        df        MS                  Number of obs = 3005
-------------+-----------------------------                F( 3, 2301) = 65.24
Model        | 50.0000000    3    18.8903724               Prob > F = 0.0000
Residual     | 650.000000 3001    .289562779               R-squared = 0.0714
-------------+-----------------------------                Adj R-squared = 0.0772
Total        | 700.000000 3004 .313782583                  Root MSE = .53811
------------------------------------------------------------------------------
lhw          | Coef.      Std. Err.     t         P>|t|    [95% Conf. Interval]
-------------+----------------------------------------------------------------
age          | .0084819   .0014144     6.00      0.000     .0057082    .0112556
female       | -.049775   .082544     -0.60      0.547    -.2116433    .1120934
femage       | -.0056949  .0020203    -2.82      0.005    -.0096568   -.0017331
_cons        | 1.762608   .0586584    30.05      0.000    1.647579     1.877637
------------------------------------------------------------------------------
```

(a) Assume the goal of the analysis is to understand the evolution of individuals' wages over time, and whether they follow a different path for females relative to male workers. Which of the two specifications above do you believe is best? Explain your reasoning.

⇒ Including tenure or not.

If we include when we should not be: SE too big → rejection problem

omitted variable problem

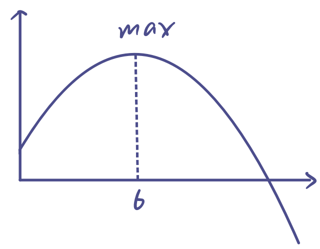
(Hypothetically, not correlated with everything ⇒ include irrelated variables)

⇒ Now its better to include . correlated ?

(b) Do wages increase, decrease or remain unchanged with every year of experience in the job? If they change, by how much? Please include a quantitative interpretation of the results, establishing the relationship between wages and tenure.

$$\frac{d\ln w}{d\,tenure} = 0.06 - 2\times 0.0005\ tenure$$

$$= 0.06 - 0.001\ tenure = 0 \quad \text{when } tenure = 6 \Rightarrow \text{maximize wage}$$



(c) Test the hypothesis that the coefficients on tenure and tenure2 are jointly significant in the model.

$$F = \frac{(RSS_{restric} - RSS_{ur})/J}{RSS_{ur}/(N-k)} \sim F_{J,N-k} \quad [RSS \text{ is } SSR]$$

$RSS_{res} = 650 \quad RSS_{ur} = 600 \quad J = 2 \quad N-k = 3005-6 = 2999$

$$F = \frac{(650-600)/2}{600/2999} = 125 > 3$$

Reject H_0: $\beta[T] = \beta[T^2] = 0$

(d) One may consider removing the variable female from the regressions. What do you expect the effect of this change to be, on the OLS estimates?

Female → uncorrelated & insignificant but you included. your variance of β will be inflated.

• Whether female is a main regressor

⇒ Small & insignificant ⇒ shouldn't including

Hypothetically if we should → remove it will cause problem of omit relevant variable. [Check later what problem will be caused because of it]

3. The following model characterizes the smoking behavior of individuals, in relation to their income:

$$Numcigs_i = \beta_0 + \beta_1 Income_i + u_i. \tag{1}$$

(a) You suspect the presence of measurement error in the left hand side (dependent) variable, which is the number of cigarettes smoked. In other words:

$$Numcigs^{\text{observed}} = Numcigs^{\text{true}} + e$$

where e is uncorrelated, simple measurement error.

Given the following information, please outline the consequences of this type of measurement error for the OLS estimation of (1). Explain your reasoning.

N = 100	$\text{Cov}\left(Numcigs, Income^{\text{true}}\right) = 5$	
	$\text{Cov}\left(Numcigs, Income^{\text{observed}}\right) = 1$	
	$\text{Var}\left(Income^{\text{true}}\right) = 5$	
	$\text{Var}(Numcigs) = 200$	
	$\text{Var}(e) = 100$	$\text{Var}(u) = 100$
	$\text{E}(u) = 0$	$\text{Cov}(e, u) = 0$
	$\text{E}(e) = 0$	

true $V(\hat{\beta}\, income) = \frac{\sigma_u^2}{N \cdot V(Income)} = \frac{100}{100 \times 5} = 0.2$

Observed $V(\hat{\beta}\, income) = \frac{\sigma_u^2 + \sigma_e^2}{N \cdot V(Income)} = \frac{100+100}{100 \times 5} = 0.4$

observed > true

True: $numcigs_i^{true} = \beta_0 + \beta_1 Income_i + u_i$

Observed: $numcigs_i^{obs} = \underbrace{\hat{\beta}_0 + \hat{\beta}_1 Income_i + u_i}_{\text{True}} + \underbrace{e_i}_{u_i}$

$\Rightarrow V(\hat{\beta}_i) = V(\underset{\text{no variance}}{\beta_1} + \frac{\frac{1}{n}\sum(X_i - \bar{X})u_i}{V(X)}) = \left[\frac{1}{n^2}\sum(X_i - \bar{X})^2\, V(u)\right] / \left[V(x)\right]^2$

$= \frac{\sigma_u^2}{n\,\cancel{n^2}} \times \cancel{n}\, \cancel{V(x)} / \cancel{V^2(x)}\ V(x)$

$= \frac{\sigma_u^2}{n V(X)}$

(b) Now consider that, instead, the right hand side variable is measured with error, as follows:

$$Income^{\text{observed}} = Income^{\text{true}} + w$$

where, again, w is simple measurement error. For this case, please find:

- the true (unobserved) β_1 capturing the relationship of income and cigarette consumption in the absence of measurement error.
- the OLS estimate that you would obtain with the data available (i.e. with the mismeasured regressor).

$$\beta^{true} = \frac{cov(x^{true}, y)}{V(x^{true})} = \frac{5}{5} = 1$$

$$\beta^{obs} = \frac{Cov(x^{obs}, y)}{V(x^{obs})} = \frac{1}{5} = 0.2$$

$$\hat{\beta}^{obs} = \frac{cov(x^{obs}, y)}{V(x^{obs})} = \beta_1 - \frac{\beta_1 V(w)}{V(x^{obs})} \neq \beta_1$$

[Measurement errors]

✓

4. True/False/Explain. To test for heteroskedastic errors in a linear model, it is useful to regress functions of the absolute values of least-squares residuals (e.g. the squared residuals) on functions of the regressors. The R-squared from this second stage regression will be (approximately) distributed as chi-square random variable under the null hypothesis of no heteroskedasticity, with degrees of freedom equal to the number of no-constant functions of the regressors in the second-stage.

good ✓ → false

$\varepsilon_i^2 = (y_i - x_i\beta)^2$ → BP test. ⇒ H_0: Homoskedasticity $\gamma_0 = 0$

H_1: $\gamma_0 \neq 0$

γ in this question is: $E[(y_i - x_i\beta)^2 | X_i] = \sigma_i^2(X) = \sigma_0^2 + Z_i\gamma_0$

if $\gamma_0 \neq 0$ ⇒ heteroskedasticity ⇒ $\sigma_i^2(X)$ change with i

⇒ Z_i: vector of independent variables from x_i. $Z_i = x_i, x_i^2$, crossproduct of x

BP test: ① OLS → $\hat{u}_i^2$

② regress $\hat{u}_i$ on Z_i + intercept ⇒ Get R^2

③ Under H_0 (homoskedasticity) $nR^2 \xrightarrow{d} \chi^2_J$ where J is the # of regressors.

5. Consider the following statement. *In applied work, it is common that the assumption of mean independence of the disturbance with respect to the matrix of regressors in the model is violated. When this is the case, OLS will not perform well as an estimator of the model parameters, but other estimators may be available.* Discuss this statement. In particular, explain what we mean by *"OLS will not perform well"*, and describe the alternative estimators that may be available. State and explain the conditions under which an alternative estimator is preferred.

F statistic ⇒ good instrument variable ⇒ F high (>10)

heteroskedasticity ⇒ F high ⇒ Good IV? Problems.

Different with different estimator.

⇒ Heterosk: OLS → GLS, FGLS, WLS, Robust

↳ Relax the assumption $E[uu'|X] = \sigma^2 I$ to $E[uu'|X] = \sigma^2 \Omega$

A4 B4

⊕ limiting behaviour of the data: $plim\left(\frac{X'\Omega X}{n}\right)$ is finite & P.d D3

[$\text{plim}\left(\frac{X'\Omega X}{n}\right) = Q$ D2]

As more data → $X'\Omega X$ bigger → $\frac{1}{n}$ by avoiding it's too big.

$\Rightarrow \text{plim}\left(\frac{X'\Omega X}{n}\right) \rightarrow$ { finite variance / off diagonal is not too large }

In GLS: OLS: unbias. consistent. asymptotically normal → not efficient usually

GLS: Ω is known $\longrightarrow \Omega = B\Lambda B' \rightarrow P = \Lambda^{-\frac{1}{2}}B' \quad P' = B\Lambda^{-\frac{1}{2}} \Rightarrow \Omega = P'P$

$y = X\beta + u \Rightarrow Py = PX\beta + \underbrace{Pu}_{\downarrow}$

check var(Pu)

$\Rightarrow Var(Pu) = E[Pu(Pu)'|PX] = P\sigma^2\Omega P' = \underbrace{\sigma^2 I}_{\downarrow}$

Classical linear regression model of Py on Px

$\hat{\beta}_{GLS} = [(PX)'(PX)]^{-1}(PX)'Py = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y$ if $\Omega = I \Rightarrow$ OLS

$Var(\hat{\beta}_{GLS}|PX) = \sigma^2((PX)'PX)^{-1} = \sigma^2(X'\Omega^{-1}X)^{-1} \Rightarrow \hat{\beta}_{GLS}$ is blue

Coming from $\min((y-X\beta)'\Omega^{-1}(y-X\beta)')$

$\Rightarrow E(Pu|PX) = 0 \longrightarrow E[\hat{\beta}_{GLS}|PX] = \beta$

↳ P is known from Ω. The requirement $E(u|X)=0$ [A3 holds]

$\Rightarrow \hat{\beta}_{GLS}$ will be consistent if $\text{plim}\left(\frac{1}{n}\underbrace{(PX)'}_{X^{*\prime}}\underbrace{PX}_{X^*}\right) = \underbrace{Q}_{Q^*}$

$\Rightarrow$ PX is well behaved. when estimator grow → PX grow

→ Sampling var: $V(\hat{\beta}|X^*) = \sigma^2(X_*'X_*)^{-1} = \underline{\sigma^2(X'\Omega^{-1}X)^{-1}}$ most efficient estimator

In FGLS: Ω is unknow & structure is know

$\hat{\Omega}$ is the estimator of $\Omega \Rightarrow \hat{\Omega} = \Omega(\hat{\theta})$ structure we know

$\hat{\beta}_{FGLS}(\hat{\Omega}) = (X'\hat{\Omega}X)^{-1}X'\hat{\Omega}^{-1}y$

How to estimate $\hat{\Omega} \Rightarrow$ need restrictions because n parameters

$\downarrow \begin{cases} \text{heteroskedasticity} \\ \text{auto correlations} \end{cases}$

N $\begin{cases} \text{finite: more efficient than LS if disturbance are non-spirical} \\ \to \infty: \hat{\beta}_{FGLS} = \hat{\beta}_{GLS} \text{ if } \operatorname{plim}\left(\frac{[X'\hat{\Omega}^{-1}X]^{-1}}{n} - \frac{(X'\Omega^{-1}X)^{-1}}{n}\right) = 0 \end{cases}$

$\Downarrow$ if $\operatorname{plim}(\hat{\theta}) = \theta$

Asymptotics $\Rightarrow$ get data from strata. bias but consistent. more data get $\to$ closer to the true value. Want $\hat{\beta}_{GLS}$. as observation growing. $\hat{\beta}_{FGLS} \to \hat{\beta}_{GLS}$

$\operatorname{plim}\left(\frac{X'\hat{\Omega}u}{\sqrt{n}} - \frac{X'\Omega^{-1}X}{\sqrt{n}}\right) = 0$

$\Rightarrow$ FGLS is the same as GLS in asymptotically.

& FGLS will be asymptotically. efficient even if $\hat{\theta}$ is not.

$\downarrow$ could be inefficient but have to be consistent

In Weighted LS: $W_{ii} = \frac{1}{w_{ii}} = \frac{1}{\sqrt{w_{ii}}} \times \frac{1}{\sqrt{w_{ii}}}$

- If Ω know, $W_{ii} = \frac{1}{w_{ii}}$ $(\sigma_i^2 = \sigma^2 x_{ik}^2)$

Special case: $Var(u|X) = E(uu'|X) = \sigma^2\Omega = \sigma^2 \begin{pmatrix} w_1 & & & 0 \\ & w_2 & & \\ & & \ddots & \\ 0 & & & w_n \end{pmatrix} \Rightarrow \Omega^{-1} = \begin{pmatrix} \frac{1}{w_1} & & & 0 \\ & \frac{1}{w_2} & & \\ & & \ddots & \\ 0 & & & \frac{1}{w_n} \end{pmatrix}$

$\Rightarrow \hat{\beta}_{WLS} = (\Sigma w_{ii} x_{ii} x_{ii}')^{-1} (\Sigma w_{ii} x_{ii} y)$ & $w_{ii} = \frac{1}{w_{ii}}$ in form of $((PX)'PX)^{-1}(PX)'y$ $\quad P = \frac{1}{\sqrt{w_{ii}}}$

More weight to the one have lower variance.

$\hat{\beta}_{WLS} \Rightarrow$ Consistent & efficient if the weights are good.

$\Rightarrow$ Consistent & Inefficient if the weights are incorrect & uncorrelated with disturbance.

- If Ω is unknown:

$\hat{\Omega} \Rightarrow$ estimator of Ω with $\hat{\Omega} = \Omega(\hat{\theta}) \Rightarrow$ To have $\hat{\Omega}$, we use a "Two step estimator"

① Estimate the model by OLS $\Rightarrow \hat{\beta}_{OLS}$ inefficient but consistent $\Rightarrow$ residual will be consistent

$\Rightarrow$ Construct the estimator of residuals $\hat{\sigma}_{ii}$

② Use these $\hat{\sigma}_{ii}$ to get $\hat{\beta}_{WLS} \Rightarrow$ more efficient. $\hat{\beta}_{WLS} = [\Sigma \frac{1}{\hat{\sigma}_i^2} x_i' x_i]^{-1} [\Sigma \frac{1}{\hat{\sigma}_i^2} x_i y_i]$

$\Rightarrow \hat{\beta}_{WLS}$ will be $\begin{cases} \text{more efficient than } \hat{\beta}_{OLS} \text{ if correct } \boxed{var(u|X)} \\ \text{less efficient than } \hat{\beta}_{OLS} \text{ if incorrect } \boxed{var(u|X)} \end{cases}$ $\sigma^2\Omega$

For robust standard error:

$\Rightarrow$ Heter u is correlated with variables in the model

$\Rightarrow \hat{AV}(\hat{\beta}) = (X'X)^{-1}(\sum \hat{u}^2 X_i' X_i)(X'X)^{-1}$

$\hat{u}_i^2$ can underestimate u_i if n is small, because OLS residuals tend to underestimate the squares of the disturbances $\rightarrow$ This is why we use $\frac{1}{n-k}$ instead of $\frac{1}{n}$ in comparing $\hat{\sigma}^2$

$\rightarrow$ rescaled by $\frac{n}{n-k}$.