Econometrics - Heteroscedasticity

Master 1 Semestre 2 - EPOLPRO (IEDES)

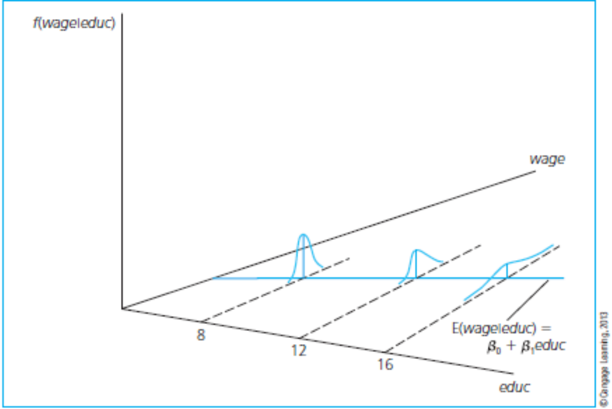
Jean-Baptiste Guiffard (Telecom-Paris, CREST)

17 février 2024

Consequences of heteroscedasticity on OLS estimates

The problem

FIGURE 2.9 Var(wage|educ) increasing with educ.



- Let's consider the following model : $y_i = \beta_0 + \beta_1 x_i + u_i$.

- Let's consider the following model : $y_i = \beta_0 + \beta_1 x_i + u_i$.
- If u_i are heteroscedastic, then $V(\widehat{\beta_1}) \neq$ constante $(\neq \frac{\sigma^2}{SCT_x})$

Consequences

No consequence on

- whether $\widehat{\beta_1}$ is biased or not; consistent or not;

Consequences on

Consequences

No consequence on

- whether $\widehat{\beta_1}$ is biased or not; consistent or not;
- whether R2 or $\overline{R2}$ are consistent or not

Consequences on

Consequences

No consequence on

- whether $\widehat{\beta_1}$ is biased or not; consistent or not;
- whether R2 or $\overline{R2}$ are consistent or not

Consequences on

- $Var(\widehat{\beta_1})$: it is biased (OLS no more efficient)

Consequences

No consequence on

- whether $\widehat{\beta_1}$ is biased or not; consistent or not;
- whether R2 or $\overline{R2}$ are consistent or not

Consequences on

- $Var(\widehat{\beta_1})$: it is biased (OLS no more efficient)
- and thus, on the validity of inferences made (based on IC, t-stat, F stat, etc.)

If u_i are heteroscedastic, then

- $V[u_i/x_i] = \sigma_i^2$

Questions are

If u_i are heteroscedastic, then

- $V[u_i/x_i] = \sigma_i^2$
- $V(\widehat{\beta}_1) = \dfrac{\Sigma(x_i - \overline{x})^2 \sigma_i^2}{(SCT_x)^2}$

Questions are

If u_i are heteroscedastic, then

- $V[u_i/x_i] = \sigma_i^2$
- $V(\widehat{\beta}_1) = \dfrac{\Sigma(x_i - \overline{x})^2\sigma_i^2}{(SCT_x)^2}$

Questions are

- Can we test the presence of heteroscedasticity ?

If u_i are heteroscedastic, then

- $V[u_i/x_i] = \sigma_i^2$
- $V(\widehat{\beta}_1) = \dfrac{\Sigma(x_i - \overline{x})^2 \sigma_i^2}{(SCT_x)^2}$

Questions are

- Can we test the presence of heteroscedasticity ?
- Is there a way to estimate $V(\widehat{\beta}_1)$ that is minimized and determine if t and F follow known distributions ?

Testing the presence of heteroscedasticity

The Breusch-Pagan test (1)

Starting point

- $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + (...) + u$ (*)

The Breusch-Pagan test (1)

Starting point

- $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + (...) + u$ (*)
- We assume (A1)' to (A4)' are verified

The Breusch-Pagan test (1)

Starting point

- $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + (...) + u$ (*)
- We assume (A1)' to (A4)' are verified
- We test H0 : $Var(u/x_1, x_2, ..., x_k) = \sigma^2$

The Breusch-Pagan test (1)

Starting point

- $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + (...) + u$ (*)
- We assume (A1)' to (A4)' are verified
- We test H0 : $Var(u/x_1, x_2, ..., x_k) = \sigma^2$
- Since $E(u/X) = 0$, then we can rewritte H0 : $Var(u/X) = E(u^2/X) = E(u^2) = \sigma^2$

The Breusch-Pagan test (1)

Starting point

- $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + (...) + u$ (*)
- We assume (A1)' to (A4)' are verified
- We test H0 : $Var(u/x_1, x_2, ..., x_k) = \sigma^2$
- Since $E(u/X) = 0$, then we can rewritte H0 : $Var(u/X) = E(u^2/X) = E(u^2) = \sigma^2$
- How test H0 ?

The Breusch-Pagan test (2)

How test H0 ?

- $u^2 = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + (...) + v$ (**)

The Breusch-Pagan test (2)

How test H0 ?

- $u^2 = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + (...) + v$ (**)
- We assume $E(v/X) = 0$

The Breusch-Pagan test (2)

How test H0 ?

- $u^2 = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + (...) + v$ (**)
- We assume $E(v/X) = 0$
- H0 : $\delta_1 = \delta_2 = \delta_k = 0$

The Breusch-Pagan test (2)

How test H0 ?

- $u^2 = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + (...) + v$ (**)
- We assume $E(v/X) = 0$
- H0 : $\delta_1 = \delta_2 = \delta_k = 0$
- In other words, testing whether u is homoscedastic in (*) amounts to testing the global significance of the model (**)

The Breusch-Pagan test (2)

How test H0 ?

- $u^2 = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + (...) + v$ (**)
- We assume $E(v/X) = 0$
- H0 : $\delta_1 = \delta_2 = \delta_k = 0$
- In other words, testing whether u is homoscedastic in (*) amounts to testing the global significance of the model (**)
- We know how to implement such a test if **() verifies the hypothesis of the CLM or if (**) verifies (A1)' to (A5)' and n is large %{so we suppose that v is homoscedastic here ?}

The Breusch-Pagan test (2)

How test H0 ?

- $u^2 = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + (...) + v$ (**)
- We assume $E(v/X) = 0$
- H0 : $\delta_1 = \delta_2 = \delta_k = 0$
- In other words, testing whether u is homoscedastic in (*) amounts to testing the global significance of the model (**)
- We know how to implement such a test if () **verifies the hypothesis of the CLM or if (**) verifies (A1)' to (A5)' and n is large %{so we suppose that v is homoscedastic here ?}
- Pb: We do not know u^2 !

The Breusch-Pagan test (3)

How test H0 ?

- $\widehat{u^2} = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + (...) + error$ (***)

The Breusch-Pagan test (3)

How test H0 ?

- $\widehat{u^2} = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + (...) + error$ (***)
- We know how to test the global significance of this last model if (***) verifies (A1)' to (A5)' and n is large {so we suppose that error is homoscedastic?}

The Breusch-Pagan test (3)

How test H0 ?

- $\widehat{u^2} = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + (...) + error$ (***)
- We know how to test the global significance of this last model if (***) verifies (A1)' to (A5)' and n is large {so we suppose that error is homoscedastic?}
- Specifically, if n is large, $F \equiv \dfrac{R^2_{\widehat{u^2}}/k}{(1 - R^2_{\widehat{u^2}})/(n-k-1)}$ follows a Fisher distribution ({not demonstrated})

The Breusch-Pagan test (3)

How test H0 ?

- $\widehat{u^2} = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + (...) + error$ (***)
- We know how to test the global significance of this last model if (***) verifies (A1)' to (A5)' and n is large {so we suppose that error is homoscedastic?}
- Specifically, if n is large, $F \equiv \frac{R^2_{\widehat{u^2}}/k}{(1 - R^2_{\widehat{u^2}})/(n-k-1)}$ follows a Fisher distribution ({not demonstrated})
  - If $F >$ critical value (read on the Fisher Table), then we reject H0 (we reject the hyp. of homoscedasticity)

The Breusch-Pagan test (3)

How test H0 ?

- $\widehat{u^2} = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + (...) + error$ (***)
- We know how to test the global significance of this last model if (***) verifies (A1)' to (A5)' and n is large {so we suppose that error is homoscedastic?}
- Specifically, if n is large, $F \equiv \frac{R^2_{\widehat{u^2}}/k}{(1 - R^2_{\widehat{u^2}})/(n-k-1)}$ follows a Fisher distribution ({not demonstrated})
  - If $F >$ critical value (read on the Fisher Table), then we reject H0 (we reject the hyp. of homoscedasticity)
  - An other version of this test is $LM \equiv n * R^2_{\widehat{u^2}}$ which follows a distribution of Khi^2_k. If $LM >$ critical value (read in the Khi^2_k Table), then we reject H0.

The Breusch-Pagan test (3)

How test H0 ?

- $\widehat{u^2} = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + (...) + error$ (***)
- We know how to test the global significance of this last model if (***) verifies (A1)' to (A5)' and n is large {so we suppose that error is homoscedastic?}
- Specifically, if n is large, $F \equiv \frac{R^2_{\widehat{u^2}}/k}{(1-R^2_{\widehat{u^2}})/(n-k-1)}$ follows a Fisher distribution ({not demonstrated})
  - If $F >$ critical value (read on the Fisher Table), then we reject H0 (we reject the hyp. of homoscedasticity)
  - An other version of this test is $LM \equiv n * R^2_{\widehat{u^2}}$ which follows a distribution of Khi^2_k. If $LM >$ critical value (read in the Khi^2_k Table), then we reject H0.
  - This last test is known as the Breusch Pagan Test

The Breusch-Pagan test (4)

Steps for implementing the Breusch-Pagan test

1. Estimate (*) and save the residuals
2. Estimate (**) and save the value of the R2
3. Compute the value of F or of LM
4. Conclude at a given significance level
5. If H0 is rejected: heteroscedasticity cannot be ignored and corrections should be applied

The White test (1)

Starting point

- When n is large, if (A1)' to (A5)' are verified, then $V(\widehat{\beta}_j)$ is unbiased and inferences based on IC, t-stat, F-stat are valid

The White test (1)

Starting point

- When n is large, if (A1)' to (A5)' are verified, then $V(\widehat{\beta}_j)$ is unbiased and inferences based on IC, t-stat, F-stat are valid
- When n is large, if (A1)' to (A4)'are verified and {there is no correlation between between the u^2 and x_j, the x_j^2 and $x_j * x_h$}, then $V(\widehat{\beta}_j)$ is unbiased and inferences based on IC, t-stat, F-stat are valid

The White test (1)

Starting point

- When n is large, if (A1)' to (A5)' are verified, then $V(\widehat{\beta}_j)$ is unbiased and inferences based on IC, t-stat, F-stat are valid
- When n is large, if (A1)' to (A4)'are verified and {there is no correlation between between the u^2 and x_j, the x_j^2 and $x_j * x_h$}, then $V(\widehat{\beta}_j)$ is unbiased and inferences based on IC, t-stat, F-stat are valid
  - In red: this weaker constraint replaces the assumption of homoscedasticity and is actually sufficient

The White test (1)

Starting point

- When n is large, if (A1)' to (A5)' are verified, then $V(\widehat{\beta_j})$ is unbiased and inferences based on IC, t-stat, F-stat are valid
- When n is large, if (A1)' to (A4)'are verified and {there is no correlation between between the u^2 and x_j, the x_j^2 and $x_j * x_h$}, then $V(\widehat{\beta_j})$ is unbiased and inferences based on IC, t-stat, F-stat are valid
  - In red: this weaker constraint replaces the assumption of homoscedasticity and is actually sufficient
- The White test is based on this result

The White test (2)

Let's write

- $\widehat{u^2} = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + \delta_3 x_1^2 + \delta_4 x_2^2 + \delta_5 x_1 x_2 + error$ $(k = 2)$ (iv)

The White test (2)

Let's write

- $\widehat{u^2} = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + \delta_3 x_1^2 + \delta_4 x_2^2 + \delta_5 x_1 x_2 + error$ $(k = 2)$ (iv)
- H0 : all δ_j (except the constant) equal zero

The White test (2)

Let's write

- $\widehat{u^2} = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + \delta_3 x_1^2 + \delta_4 x_2^2 + \delta_5 x_1 x_2 + error$ $(k = 2)$ (iv)
- H0 : all δ_j (except the constant) equal zero
- We know how to implement this test

The White test (2)

Let's write

- $\widehat{u^2} = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + \delta_3 x_1^2 + \delta_4 x_2^2 + \delta_5 x_1 x_2 + error$ $(k = 2)$ (iv)
- H0 : all δ_j (except the constant) equal zero
- We know how to implement this test
- Drawbacks ?

The White test (2)

Let's write

- $\widehat{u^2} = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + \delta_3 x_1^2 + \delta_4 x_2^2 + \delta_5 x_1 x_2 + error$ $(k = 2)$ (iv)
- H0 : all δ_j (except the constant) equal zero
- We know how to implement this test
- Drawbacks ?
  - If at a start $k = 2$, # of paramters to estimate for the test amounts to 5; If at a start $k = 3$, # of paramters to estimate for the test amounts to 9; If at a start $k = 6$, # of paramters to estimate for the test amounts to 27;

The White test (2)

Let's write

- $\widehat{u^2} = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + \delta_3 x_1^2 + \delta_4 x_2^2 + \delta_5 x_1 x_2 + error$ $(k = 2)$ (iv)
- H0 : all δ_j (except the constant) equal zero
- We know how to implement this test
- Drawbacks ?
  - If at a start $k = 2$, # of paramters to estimate for the test amounts to 5; If at a start $k = 3$, # of paramters to estimate for the test amounts to 9; If at a start $k = 6$, # of paramters to estimate for the test amounts to 27;
  - Could be very demanding in *ddl* !

The White test (2)

Let's write

- $\widehat{u^2} = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + \delta_3 x_1^2 + \delta_4 x_2^2 + \delta_5 x_1 x_2 + error$ $(k = 2)$ (iv)
- H0 : all δ_j (except the constant) equal zero
- We know how to implement this test
- Drawbacks ?
    - If at a start $k = 2$, # of paramters to estimate for the test amounts to 5; If at a start $k = 3$, # of paramters to estimate for the test amounts to 9; If at a start $k = 6$, # of paramters to estimate for the test amounts to 27;
    - Could be very demanding in *ddl* !
    - An alternative is to synthetize the information using $\widehat{y}$ and $\widehat{y}^2$ instead of the X

The White test (2)

Let's write

- $\widehat{u^2} = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + \delta_3 x_1^2 + \delta_4 x_2^2 + \delta_5 x_1 x_2 + error$ $(k = 2)$ (iv)
- H0 : all δ_j (except the constant) equal zero
- We know how to implement this test
- Drawbacks ?
  - If at a start $k = 2$, # of paramters to estimate for the test amounts to 5; If at a start $k = 3$, # of paramters to estimate for the test amounts to 9; If at a start $k = 6$, # of paramters to estimate for the test amounts to 27;
  - Could be very demanding in *ddl* !
  - An alternative is to synthetize the information using $\widehat{y}$ and $\widehat{y}^2$ instead of the X
  - $\widehat{u^2} = \gamma_0 + \gamma_1 \widehat{y} + \gamma_2 \widehat{y}^2 + error$ (v)

The White test (2)

Let's write

- $\widehat{u^2} = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + \delta_3 x_1^2 + \delta_4 x_2^2 + \delta_5 x_1 x_2 + error$ $(k = 2)$ (iv)
- H0 : all δ_j (except the constant) equal zero
- We know how to implement this test
- Drawbacks ?
    - If at a start $k = 2$, # of paramters to estimate for the test amounts to 5; If at a start $k = 3$, # of paramters to estimate for the test amounts to 9; If at a start $k = 6$, # of paramters to estimate for the test amounts to 27;
    - Could be very demanding in *ddl* !
    - An alternative is to synthetize the information using $\widehat{y}$ and $\widehat{y}^2$ instead of the X
    - $\widehat{u^2} = \gamma_0 + \gamma_1 \widehat{y} + \gamma_2 \widehat{y}^2 + error$ (v)
    - H0: $\gamma_1 = \gamma_2 = 0$

The White test (3)

Steps for implementing the White test

1. Estimate (*) and save the residuals and the $\widehat{y}$
2. Estimate (v) and save the value of the R2
3. Compute the value of F (or of LM)
4. Conclude at a given significance level
5. If H0 is rejected: heteroscedasticity cannot be ignored and corrections should be applied

Let's suppose H0 is rejected. What do we do then ?

- Option 1: We estimate the model in OLS BUT we correct for $V(\widehat{\beta}_j)$

- Option 1: We estimate the model in OLS BUT we correct for $V(\widehat{\beta}_j)$
- Option 2: We estimate the model using weighted least square

- Option 1: We estimate the model in OLS BUT we correct for $V(\widehat{\beta}_j)$
- Option 2: We estimate the model using weighted least square
- Both methods have drawbacks and advantages !

OLS inference robust to heteroscedasticity

White robust standard errors

- We consider the model: $y = \beta_0 + \beta_1 x_1 + u$

White robust standard errors

- We consider the model: $y = \beta_0 + \beta_1 x_1 + u$
- We assume (A1)' to (A4)' are verified

White robust standard errors

- We consider the model: $y = \beta_0 + \beta_1 x_1 + u$
- We assume (A1)' to (A4)' are verified
  - Thus, $\widehat{\beta_1} = \dfrac{\Sigma(x-\overline{x})(y-\overline{y})}{\Sigma(x-\overline{x})^2}$ or $= \beta_1 + \dfrac{\Sigma(x-\overline{x})u}{\Sigma(x-\overline{x})^2}$

White robust standard errors

- We consider the model: $y = \beta_0 + \beta_1 x_1 + u$
- We assume (A1)' to (A4)' are verified
  - Thus, $\widehat{\beta_1} = \frac{\Sigma(x-\overline{x})(y-\overline{y})}{\Sigma(x-\overline{x})^2}$ or $= \beta_1 + \frac{\Sigma(x-\overline{x})u}{\Sigma(x-\overline{x})^2}$
- u are heteroscedastic

White robust standard errors

- We consider the model: $y = \beta_0 + \beta_1 x_1 + u$
- We assume (A1)' to (A4)' are verified
  - Thus, $\widehat{\beta_1} = \frac{\Sigma(x-\overline{x})(y-\overline{y})}{\Sigma(x-\overline{x})^2}$ or $= \beta_1 + \frac{\Sigma(x-\overline{x})u}{\Sigma(x-\overline{x})^2}$
- u are heteroscedastic
  - Then, $V(\widehat{\beta_1})$ does not simplify to equal $\frac{\sigma^2}{SCT_x}$

White robust standard errors

- We consider the model: $y = \beta_0 + \beta_1 x_1 + u$
- We assume (A1)' to (A4)' are verified
  - Thus, $\widehat{\beta_1} = \frac{\Sigma(x-\overline{x})(y-\overline{y})}{\Sigma(x-\overline{x})^2}$ or $= \beta_1 + \frac{\Sigma(x-\overline{x})u}{\Sigma(x-\overline{x})^2}$
- u are heteroscedastic
  - Then, $V(\widehat{\beta_1})$ does not simplify to equal $\frac{\sigma^2}{SCT_x}$
  - $V(\widehat{\beta_1}) = \frac{\Sigma(SCT_x * V[u/x])}{SCT_x^2}$

White robust standard errors

- We consider the model: $y = \beta_0 + \beta_1 x_1 + u$
- We assume (A1)' to (A4)' are verified
  - Thus, $\widehat{\beta_1} = \frac{\Sigma(x-\overline{x})(y-\overline{y})}{\Sigma(x-\overline{x})^2}$ or $= \beta_1 + \frac{\Sigma(x-\overline{x})u}{\Sigma(x-\overline{x})^2}$
- u are heteroscedastic
  - Then, $V(\widehat{\beta_1})$ does not simplify to equal $\frac{\sigma^2}{SCT_x}$
  - $V(\widehat{\beta_1}) = \frac{\Sigma(SCT_x * V[u/x])}{SCT_x^2}$
  - How estimate $V(\widehat{\beta_1})$?

White robust standard errors

- We consider the model: $y = \beta_0 + \beta_1 x_1 + u$
- We assume (A1)' to (A4)' are verified
  - Thus, $\widehat{\beta_1} = \dfrac{\Sigma(x-\overline{x})(y-\overline{y})}{\Sigma(x-\overline{x})^2}$ or $= \beta_1 + \dfrac{\Sigma(x-\overline{x})u}{\Sigma(x-\overline{x})^2}$
- u are heteroscedastic
  - Then, $V(\widehat{\beta_1})$ does not simplify to equal $\dfrac{\sigma^2}{SCT_x}$
  - $V(\widehat{\beta_1}) = \dfrac{\Sigma(SCT_x * V[u/x])}{SCT_x^2}$
  - How estimate $V(\widehat{\beta_1})$?
- White (1980) suggests to measure $\widehat{V(\widehat{\beta_1})}$ with $\dfrac{\Sigma(SCT_x * \widehat{u}^2)}{SCT_x^2}$ {(not demonstrated – for a discussion see p396 in Wooldridge (2013))}.

White robust standard errors (2)

- The corresponding standard errors are named White robust standard errors.

White robust standard errors (2)

- The corresponding standard errors are named White robust standard errors.
- If n is large, and if we use the White robust standard errors, the t-stat (respectively, the F-stat, named Wald stat) follow a Student distribution (respectively, a Fisher distribution) $\rightarrow$ inference is valid

White robust standard errors (2)

- The corresponding standard errors are named White robust standard errors.
- If n is large, and if we use the White robust standard errors, the t-stat (respectively, the F-stat, named Wald stat) follow a Student distribution (respectively, a Fisher distribution) $\rightarrow$ inference is valid
- If n is *NOT* large, and if we use the White robust standard errors, the t-stat may not follow a Student distribution $\rightarrow$ inference is not valid

White robust standard errors (2)

- The corresponding standard errors are named White robust standard errors.
- If n is large, and if we use the White robust standard errors, the t-stat (respectively, the F-stat, named Wald stat) follow a Student distribution (respectively, a Fisher distribution) $\rightarrow$ inference is valid
- If n is *NOT* large, and if we use the White robust standard errors, the t-stat may not follow a Student distribution $\rightarrow$ inference is not valid
- [$\rightarrow$] If n large: report systematically the robust standard errors and perform the usual tests, as usual (the command is automatized in most statistical softwares)

White robust standard errors (2)

- The corresponding standard errors are named White robust standard errors.
- If n is large, and if we use the White robust standard errors, the t-stat (respectively, the F-stat, named Wald stat) follow a Student distribution (respectively, a Fisher distribution) $\rightarrow$ inference is valid
- If n is *NOT* large, and if we use the White robust standard errors, the t-stat may not follow a Student distribution $\rightarrow$ inference is not valid
- [$\rightarrow$] If n large: report systematically the robust standard errors and perform the usual tests, as usual (the command is automatized in most statistical softwares)
- [$\rightarrow$] If n *NOT* large, prefer alternative methods

Weighted least square estimation

The WLS estimator (1)

Starting point

- $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + (...) + u$ (*)

The WLS estimator (1)

Starting point

- $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + (...) + u$ (*)
- We assume (A1)' to (A4)' are verified

The WLS estimator (1)

Starting point

- $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + (...) + u$ (*)
- We assume (A1)' to (A4)' are verified
- We also assume that $V(u/X) = E(u^2/X) = E(u^2) = \sigma^2 * h(X)$ with $h(x) > 0$ (since V()>0) ($h(X)$ is assumed to be known)

The WLS estimator (2)

- We show that

The WLS estimator (2)

- We show that
  - $E(\frac{u}{\sqrt{h}}/X) = 0$

The WLS estimator (2)

- We show that
  - $E(\frac{u}{\sqrt{h}}/X) = 0$
  - $V(\frac{u}{\sqrt{h}}/X) = E((\frac{u}{\sqrt{h}})^2/X) = E((\frac{u}{\sqrt{h}})^2) = \frac{E(u^2)}{h} = \frac{\sigma^2 * h}{h} = \sigma^2$

The WLS estimator (2)

- We show that
  - $E(\frac{u}{\sqrt{h}}/X) = 0$
  - $V(\frac{u}{\sqrt{h}}/X) = E((\frac{u}{\sqrt{h}})^2/X) = E((\frac{u}{\sqrt{h}})^2) = \frac{E(u^2)}{h} = \frac{\sigma^2 * h}{h} = \sigma^2$
- Let's rewritte (*) : $\frac{y}{\sqrt{h}} = \frac{\beta_0}{\sqrt{h}} + \frac{\beta_1}{\sqrt{h}}x_1 + \frac{\beta_2}{\sqrt{h}}x_2 + (...) + \frac{u}{\sqrt{h}}$ (++)

The WLS estimator (2)

- We show that
  - $E(\frac{u}{\sqrt{h}}/X) = 0$
  - $V(\frac{u}{\sqrt{h}}/X) = E((\frac{u}{\sqrt{h}})^2/X) = E((\frac{u}{\sqrt{h}})^2) = \frac{E(u^2)}{h} = \frac{\sigma^2 * h}{h} = \sigma^2$
- Let's rewritte (*) : $\frac{y}{\sqrt{h}} = \frac{\beta_0}{\sqrt{h}} + \frac{\beta_1}{\sqrt{h}}x_1 + \frac{\beta_2}{\sqrt{h}}x_2 + (...) + \frac{u}{\sqrt{h}}$ (++)
  - (A1)' to (A4)' are still verified

The WLS estimator (2)

- We show that
  - $E(\frac{u}{\sqrt{h}}/X) = 0$
  - $V(\frac{u}{\sqrt{h}}/X) = E((\frac{u}{\sqrt{h}})^2/X) = E((\frac{u}{\sqrt{h}})^2) = \frac{E(u^2)}{h} = \frac{\sigma^2 * h}{h} = \sigma^2$
- Let's rewritte (*) : $\frac{y}{\sqrt{h}} = \frac{\beta_0}{\sqrt{h}} + \frac{\beta_1}{\sqrt{h}}x_1 + \frac{\beta_2}{\sqrt{h}}x_2 + (...) + \frac{u}{\sqrt{h}}$ (++)
  - (A1)' to (A4)' are still verified
  - In this last model: $V(\frac{u}{\sqrt{h}}/X) = \sigma^2$ i.e. (A5)' is also verified

The WLS estimator (2)

- We show that
  - $E(\frac{u}{\sqrt{h}}/X) = 0$
  - $V(\frac{u}{\sqrt{h}}/X) = E((\frac{u}{\sqrt{h}})^2/X) = E((\frac{u}{\sqrt{h}})^2) = \frac{E(u^2)}{h} = \frac{\sigma^2 * h}{h} = \sigma^2$
- Let's rewritte (*) : $\frac{y}{\sqrt{h}} = \frac{\beta_0}{\sqrt{h}} + \frac{\beta_1}{\sqrt{h}}x_1 + \frac{\beta_2}{\sqrt{h}}x_2 + (...) + \frac{u}{\sqrt{h}}$ (++)
  - (A1)' to (A4)' are still verified
  - In this last model: $V(\frac{u}{\sqrt{h}}/X) = \sigma^2$ i.e. (A5)' is also verified
  - If we further assume that u in (*) is normal, then $\frac{u}{\sqrt{h}}$ is normal, then (A6)' is also verified

The WLS estimator (2)

- We show that
  - $E(\frac{u}{\sqrt{h}}/X) = 0$
  - $V(\frac{u}{\sqrt{h}}/X) = E((\frac{u}{\sqrt{h}})^2/X) = E((\frac{u}{\sqrt{h}})^2) = \frac{E(u^2)}{h} = \frac{\sigma^2 * h}{h} = \sigma^2$
- Let's rewritte (*) : $\frac{y}{\sqrt{h}} = \frac{\beta_0}{\sqrt{h}} + \frac{\beta_1}{\sqrt{h}}x_1 + \frac{\beta_2}{\sqrt{h}}x_2 + (...) + \frac{u}{\sqrt{h}}$ (++)
  - (A1)' to (A4)' are still verified
  - In this last model: $V(\frac{u}{\sqrt{h}}/X) = \sigma^2$ i.e. (A5)' is also verified
  - If we further assume that u in (*) is normal, then $\frac{u}{\sqrt{h}}$ is normal, then (A6)' is also verified
- The hypothesis of the CLM being verified in (++), we can estimate it in OLS and obtain unbiased $\widehat{\beta}_j^*$ (the coefficient in (++)) and $V(\widehat{\beta}_j^*)$

The WLS estimator (2)

- We show that
  - $E(\frac{u}{\sqrt{h}}/X) = 0$
  - $V(\frac{u}{\sqrt{h}}/X) = E((\frac{u}{\sqrt{h}})^2/X) = E((\frac{u}{\sqrt{h}})^2) = \frac{E(u^2)}{h} = \frac{\sigma^2 * h}{h} = \sigma^2$
- Let's rewritte (*) : $\frac{y}{\sqrt{h}} = \frac{\beta_0}{\sqrt{h}} + \frac{\beta_1}{\sqrt{h}}x_1 + \frac{\beta_2}{\sqrt{h}}x_2 + (...) + \frac{u}{\sqrt{h}}$ $(++)$
  - (A1)' to (A4)' are still verified
  - In this last model: $V(\frac{u}{\sqrt{h}}/X) = \sigma^2$ i.e. (A5)' is also verified
  - If we further assume that u in (*) is normal, then $\frac{u}{\sqrt{h}}$ is normal, then (A6)' is also verified
- The hypothesis of the CLM being verified in $(++)$, we can estimate it in OLS and obtain unbiased $\widehat{\beta}_j^*$ (the coefficient in $(++)$) and $V(\widehat{\beta}_j^*)$
- Recall $V(\widehat{\beta}_j)$ are biased $\rightarrow$ prefer estimating $(++)$

The WLS estimator (3)

- $\widehat{\beta}_j^*$ are named weighted least square estimates; there are an exemple of what we call generalized least square estimates

The WLS estimator (3)

- $\widehat{\beta}_j^*$ are named weighted least square estimates; there are an exemple of what we call generalized least square estimates
- They inherit the interpretation of $\widehat{\beta}_j$

The WLS estimator (3)

- $\widehat{\beta}_j^*$ are named weighted least square estimates; there are an exemple of what we call generalized least square estimates
- They inherit the interpretation of $\widehat{\beta}_j$
- However the $R2$ in $(++)$ is not informative for (*)

The WLS estimator (4)

- The estimation of $(++)$ is automatized in most statistical softwares; we just need to specify h

The WLS estimator (4)

- The estimation of $(++)$ is automatized in most statistical softwares; we just need to specify h
- In most cases, the definition of h remains arbitrary. What if h is misspecified ? Still better to estimate $(++)$ than (*)

The WLS estimator (4)

- The estimation of (++) is automatized in most statistical softwares; we just need to specify h
- In most cases, the definition of h remains arbitrary. What if h is misspecified ? Still better to estimate (++) than (*)
  - Apply the white correction method on (++) ?