Lecture 05

White's test

- Does not rely upon a specific form of heteroskedasticity.
- It's closely related to the model: $\sigma_i^2 = \alpha_1 + \alpha_2 z_{2i} + \ldots + \alpha_p z_{pi}$;
  - It's a large sample LM test;
- Does not require normality of the errors.

Example

$$y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + u_i, \quad \mathrm{Var}(u_i) = \sigma_i^2$$

Model of the variance:

$$\sigma_i^2 = \alpha_1 + \alpha_2 x_{2i} + \alpha_3 x_{3i} + \alpha_4 x_{2i}^2 + \alpha_5 x_{3i}^2 + \alpha_6 x_{2i} x_{3i}$$

- All variables, all their squares and all the (unique) inner products.

$$\begin{aligned} H_0 &: \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = \alpha_6 = 0 \\ H_1 &: \exists\, \alpha_j \neq 0 \quad (j \in \{2,3,4,5,6\}) \end{aligned}$$

- Steps:
  1. Estimate the original model by OLS;
  2. Find the residuals: $\hat{u}_i = y_i - \hat{\beta}_1 - \hat{\beta}_2 x_{2i} - \hat{\beta}_3 x_{3i}$;
  3. Regress the variance model, using $\hat{u}_i$ as an estimator for σ_i^2 ;
  4. Find the test statistic NR^2 which follows a χ^2 distribution:
     - $NR^2 \sim \chi^2(p-1)$, where $p-1$ is the number of restrictions (in the example, it would be 5, after the five α's in H_0);
  5. Compare to the critical value for the given level of significance (for instance, $\chi^2_{0.05}(5) = 11.07$):
     - If the test statistic is greater than the critical value, reject H_0;
     - Otherwise, we fail to reject H_0.

Estimation

1. White's correction:
   - White's Heteroskedasticity-Consistent Covariance Matrix (HCCM)
   - Non-parametric test that re-weights the variable, regarding the variance (since OLS gives the variables with more variance more weight).
2. Generalized (or Weighted) Least Squares
   - Consists in re-weighting the observations.

Example

Weighted Least Squares

Linear model: $y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + u_i$;

Variance model: $\sigma_i^2 = \mathrm{Var\ for\ error}\ i$.

Dividing both sides of the linear model by σ_i:

$$\frac{y_i}{\sigma_i} = \beta_1 \frac{1}{\sigma_i} + \beta_2 \frac{x_{2i}}{\sigma_i} + \beta_3 \frac{x_{3i}}{\sigma_i} + \frac{u_i}{\sigma_i}$$

Renaming it:

$$y_i^* = \beta_1 x_{1i}^* + \beta_2 x_{2i}^* + \beta_3 x_{3i}^* + u_i^*$$

- Because of heteroskedasticity, the OLS is not BLUE for the y_i model; it is BLUE, however, for the y_i^* model, in which there is no heteroskedasticity.

Why is it?

- Dividing both sides of the model by σ_i up-weights the observations with low variance (which would otherwise be underrated by OLS) and down-weights those with high variance (overrated by OLS).
- Also, it was defined that $\text{Var}(u_i^*) := \text{Var}\left(\frac{u_i}{\sigma_i}\right)$.
  - A property of the variance is $\text{Var}(a + bX) = b^2 + \text{Var}(X)$, where a, b are constants and X, a random variable.
  - Then, $\text{Var}(u_i^*) := \text{Var}\left(\frac{u_i}{\sigma_i}\right) = \frac{1}{\sigma_i^2}\text{Var}(u_i)$, since σ_i^2 is a number (constant).
  - Given that $\text{Var}(u_i) = \sigma_i^2$, then $\text{Var}(u_i^*) = \frac{1}{\sigma_i^2}\sigma_i^2 = 1$.
  - With $\text{Var}(u_i^*)$ being constant, the y_i^* model features homoskedasticity.
- Conclusion: by re-weighting the variables (dividing them by σ_i), it is possible to get rid of heteroskedasticity - which makes OLS BLUE -, provided that σ_i is known.
  - It is possible to use the LM models for variance from Lecture 03 (*Breusch-Pagan*, *Glejser* or *Park*) to estimate σ_i and make the procedure work.

Matrix form

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{u}$$

where $\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_N \end{bmatrix}$, $\boldsymbol{X} = \begin{bmatrix} 1 & x_{21} & x_{31} & \cdots & x_{p1} \\ 1 & x_{22} & x_{32} & \cdots & x_{p2} \\ 1 & x_{23} & x_{33} & \cdots & x_{p3} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & x_{2N} & x_{3N} & \cdots & x_{pN} \end{bmatrix}$, $\boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \\ \beta_p \end{bmatrix}$, $\boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \\ u_N \end{bmatrix}$.

Then, the estimator is given by $\hat{\boldsymbol{\beta}}_{OLS} = \left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1}\boldsymbol{X}^T\boldsymbol{y}$ (normal equation).

- This is the vectorized, multivariate version for $\beta = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2}$.

$\Omega = \text{Var}(\boldsymbol{u})$ is the variance-covariance matrix.

- $\Omega = \boldsymbol{u}\boldsymbol{u}^T = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_N \end{bmatrix} \begin{bmatrix} u_1 & u_2 & \cdots & u_N \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_N^2 \end{bmatrix}$.
- All the elements of Ω are inner products $u_i u_j$, so that $i, j \in \{1, 2, \ldots, N\}$.
  - Those in the main diagonal are the variances for each observation.

- If the errors are uncorrelated, every $u_i u_j$ where $i \neq j$ (that is, all elements outside the main diagonal) are equal to zero.
- If there is homoskedasticity, then all the element in the main diagonal are the same (*i. e.*, $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_N^2$).

> "So, the problems you can run into are twofold: either the elements on the diagonal of your variance-covariance matrix are not the same (**heteroskedasticity**), or there is correlation among your errors, which is what we call **autocorrelation**."

- Ω is a positive-definite matrix (more on PDMs here):
  - As every positive-definite matrix, it can be decomposed in the form $\Omega = Q\Lambda Q^T$ - where the columns of Q are the eigenvectors of Ω and the diagonal entries of matrix Λ, the respective eigenvalues (more here).
  - Under that condition, it is possible to write $\Omega = QIQ^T$ (where I is the identity matrix), without loss of generality.
  - A correction technique consists in transforming the model so that, instead of having Ω as the variance-covariance matrix, we have I .
    - $Q^{-1}\Omega(Q^{-1})^T = I \quad \Rightarrow \quad Q^{-1}\boldsymbol{u}\boldsymbol{u}^T(Q^{-1})^T = I$.
    - $E\left(\boldsymbol{u}\boldsymbol{u}^T\right) = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_N^2 \end{bmatrix}$.
    - That can be decomposed $E\left(\boldsymbol{u}\boldsymbol{u}^T\right) = QIQ^T$, where

$$Q = Q^T = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_N \end{bmatrix}. \text{ Thus, } Q^{-1} = \begin{bmatrix} \frac{1}{\sigma_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sigma_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sigma_N} \end{bmatrix}.$$

- Then, the model is re-weighted: $Q^{-1}\boldsymbol{y} = Q^{-1}\left(\boldsymbol{X\beta} + \boldsymbol{u}\right)$.
  - In the case shown in the previous section (WLS), the diagonal of Q would formed by the variances (with zero elsewhere). Consequently, the diagonal of Q^{-1} would be filled by the reciprocals of those variances: $\left(\sigma_i^2\right)^{-1}$.
  - Then, pre-multiplying both sides of the model by Q^{-1} has the same effect as dividing both sides by the variances.

- The form $\boldsymbol{\beta} = \left(\boldsymbol{X}^T\Omega^{-1}\boldsymbol{X}\right)^{-1}\left(\boldsymbol{X}^T\Omega^{-1}\boldsymbol{y}\right)$ - derived from the normal equation - is called **Generalized Least Squares** (GLS) estimator.
  - If there is no heteroskedasticity nor autocorrelation, $\Omega = I$ and the GLS is equal to the normal equation.
  - Otherwise, the model is re-weighted by the variance-covariance matrix.

Model: $y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + u_i$, $\mathrm{Var}(u_i) = \sigma_i^2$.

Method: Re-weight by $\dfrac{1}{\sigma_1}$.

Problem: We do not know σ_i .

Solution: Use a model of variance as a way to estimate σ_i (that is, use $\hat{\sigma}_i$ instead of σ_i) .

- Simplest case: $\mathrm{Var}(u_i) = \sigma^2 Z_i^2 \equiv \sigma_i^2$, where Z_i^2 is known data.

- Then, $\sigma_i = \sigma Z_i$;
- We divide the model by Z_i : $\frac{y_i}{Z_i} = \beta_1 \frac{1}{Z_i} + \beta_2 \frac{x_{2i}}{Z_i} + \beta_3 \frac{x_{3i}}{Z_i} + \frac{u_i}{Z_i}$.
- Then, we re-write it: $y_i^* = \beta_1 x_{1i}^* + \beta_2 x_{2i}^* + \beta_3 x_{3i}^* + u_i^*$.
- $\mathrm{Var}(u_i^*) = \mathrm{Var}\left(\frac{u_i}{Z_i}\right) = \frac{1}{Z_i^2}\mathrm{Var}(u_i) = \frac{1}{Z_i^2}\left(\sigma^2 Z_i^2\right) = \sigma^2$ (homoskedasticity) .

Feasible Generalized Least Squares (FGLS)

Variance models:

(a) $\sigma_i^2 = \alpha_1 + \alpha_2 z_{2i} + \ldots + \alpha_p z_{pi}$;

(b) $\sigma_i = \alpha_1 + \alpha_2 z_{2i} + \ldots + \alpha_p z_{pi}$;

(c) $\ln \sigma_i^2 = \alpha_1 + \alpha_2 z_{2i} + \ldots + \alpha_p z_{pi}$.

Steps:

1. Regress y on constant, $x_2, \ldots, x_k$: get $\hat{\beta}_{OLS}$;
2. Compute $\hat{u}_i = y_i - \hat{\beta}_1 - \hat{\beta}_2 x_{2i} - \ldots - \hat{\beta}_k x_{ki}$;
3. (a) Regress $\hat{u}^2$ on constant, $z_2, \ldots, z_p$: get $\hat{\alpha}$ and apply to $\sigma_i^2 = \hat{\alpha}_1 + \hat{\alpha}_2 z_{2i} + \ldots + \hat{\alpha}_p z_{pi}$;
4. Divide the linear model by $\hat{\sigma}_i$: $\frac{y_i}{\hat{\sigma}_i} = \beta_1 \frac{1}{\hat{\sigma}_i} + \beta_2 \frac{x_{2i}}{\hat{\sigma}_i} + \ldots + \beta_k \frac{x_{ki}}{\hat{\sigma}_i} + \frac{u_i}{\hat{\sigma}_i}$.

Problems:

- There is no guarantee that $\hat{\sigma}_i^2 > 0$: use absolute value.