Lecture 06

Steps for correcting heteroskedasticity

1. Regress y on constant, $x_2, \dots, x_k$;
   - This produces $\hat{\beta}_{OLS}$.
2. Calculate the estimated error (residuals): $\hat{u}_i = y_i - \hat{\beta}_1 - \hat{\beta}_2 x_{2i} - \dots - \hat{\beta}_k x_{ki}$;
3. (a) Regress $\hat{u}_i^2$ on constant, $z_2, \dots, z_p$;
   - $\hat{u}_i^2 = \alpha_1 + \alpha_2 z_2 + \dots \alpha_p z_p + \text{error}$;
   - This produces $\hat{\alpha}_i$.
4. Use the predicted variance $\hat{\sigma}_i^2$ to get $\hat{\sigma}_i$:
   - $\hat{\sigma}_i^2 = \hat{\alpha}_1 + \hat{\alpha}_2 z_2 + \dots + \hat{\alpha}_p z_p \quad \Rightarrow \quad \hat{\sigma}_i^2 = \hat{u}_i^2 - \text{error}$;
   - Since there is no guarantee that $\hat{\sigma}_i^2 > 0$, it might be a good idea to use the absolute value.
5. Divide the whole original model by $\hat{\sigma}_i$:
   - $\dfrac{y_i}{\hat{\sigma}_i} = \beta_1 \dfrac{1}{\hat{\sigma}_i} + \beta_2 \dfrac{x_{2i}}{\hat{\sigma}_i} + \dots + \beta_k \dfrac{x_{ki}}{\hat{\sigma}_i} + \dfrac{u_i}{\hat{\sigma}_i}$.
6. Obtain the new model: $y^* = \beta_1 + \beta_2 x_2^* + \dots + \beta_k x_k^* + u^*$;
   - This model features no heteroskedasticity.
   - OLS is BLUE.

Suppose you have **serial correlation** (autocorrelation) and you do OLS:

1. Coefficients remain unbiased and consistent, but not if there is a lagged dependent variable.

$$y_t = \beta_1 + \beta_2 y_{t-1} + \beta_3 x_t + u_t, \quad t = 1, 2, \dots, T$$

2. OLS is inefficient: better estimator exists.
3. The standard errors of β_i are wrong:
   - In Economics, they are often biased downwards, when there is positive serial correlation.
   - In this case, the t statistic is overestimated, since $t = \dfrac{\beta^{H_0} - \hat{\beta}}{\hat{\sigma}}$.

Generalization of assumptions for an estimator to be BLUE:

1. The model is linear;
2. The x's are non-stochastic;
3. There is no perfect multicollinearity;
4. The error has zero mean;
5. The errors are homoskedastic;
6. $u_t \perp u_s, \ \forall \, t \neq s$ (errors are independent):
   - Common model: $u_t = \rho u_{t-1} + e_t, \quad |\rho| < 1$.
7. The u's and the x's are uncorrelated:

- If $y_t = \beta_1 + \beta_2 y_{t-1} + \beta_3 x_t + u_t$ and $u_t = \rho u_{t-1} + e_t$, then y_t can be rewritten as $y_t = \beta_1 + \beta_2 y_{t-1} + \beta_3 x_t + (\rho u_{t-1} + e_t)$;
- There is a clear correlation between y_{t-1} and u_{t-1}.

8. The errors are normally distributed.