Heteroskedasticity and Autocorrelation

Fall 2008

Heteroskedasticity

- We now relax the assumption of homoskedasticity, while all other assumptions remain to hold.
- Heteroskedasticity is said to occur when the variance of the unobservable error u, conditional on independent variables, is not constant.

$$Var\left(u_i | X_i\right) = \sigma_i^2.$$

- In particular, the variance of the error may be a function of independent variables:

$$Var\left(u_i | X_i\right) = \sigma^2 h\left(X_i\right).$$

Consequences of Heteroskedasticity

- First, note that we do not need the homoskedasticity asssumption to show the unbiasedness of OLS. Thus, *OLS is still unbiased*.

Consequences of Heteroskedasticity

- First, note that we do not need the homoskedasticity asssumption to show the unbiasedness of OLS. Thus, *OLS is still unbiased*.
- However, the homoskedasticity assumption is needed to show the efficiency of OLS. Hence, *OLS is not BLUE any longer*.

Consequences of Heteroskedasticity

- First, note that we do not need the homoskedasticity asssumption to show the unbiasedness of OLS. Thus, *OLS is still unbiased*.
- However, the homoskedasticity assumption is needed to show the efficiency of OLS. Hence, *OLS is not BLUE any longer*.
- The variances of the OLS estimators are biased in this case. Thus, the usual OLS t statistic and confidence intervals are no longer valid for inference problem.

Consequences of Heteroskedasticity

- First, note that we do not need the homoskedasticity asssumption to show the unbiasedness of OLS. Thus, *OLS is still unbiased*.
- However, the homoskedasticity assumption is needed to show the efficiency of OLS. Hence, *OLS is not BLUE any longer*.
- The variances of the OLS estimators are biased in this case. Thus, the usual OLS t statistic and confidence intervals are no longer valid for inference problem.
- We can still use the OLS estimators by finding *heteroskedasticity-robust estimators* of the variances.

Consequences of Heteroskedasticity

- First, note that we do not need the homoskedasticity asssumption to show the unbiasedness of OLS. Thus, *OLS is still unbiased*.
- However, the homoskedasticity assumption is needed to show the efficiency of OLS. Hence, *OLS is not BLUE any longer*.
- The variances of the OLS estimators are biased in this case. Thus, the usual OLS t statistic and confidence intervals are no longer valid for inference problem.
- We can still use the OLS estimators by finding *heteroskedasticity-robust estimators* of the variances.
- Alternatively, we can devise an *efficient* estimator by re-weighting the data appropriately to take into account of heteroskedasticity.

How Does the Heteroskedasticity Look?

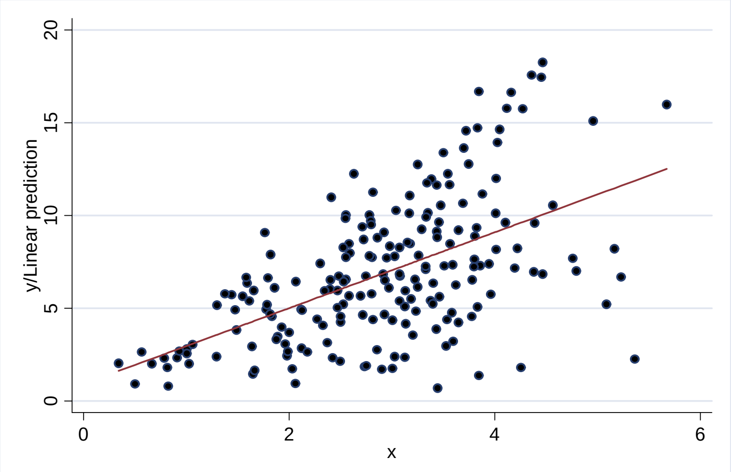
- Consider the following true regression model with heteroskedastic errors:

$$Y_i = 1 + 2X_i + u_i, \text{ where } u_i \sim N\left(0, X_i^2\right)$$

How Does the Heteroskedasticity Look?

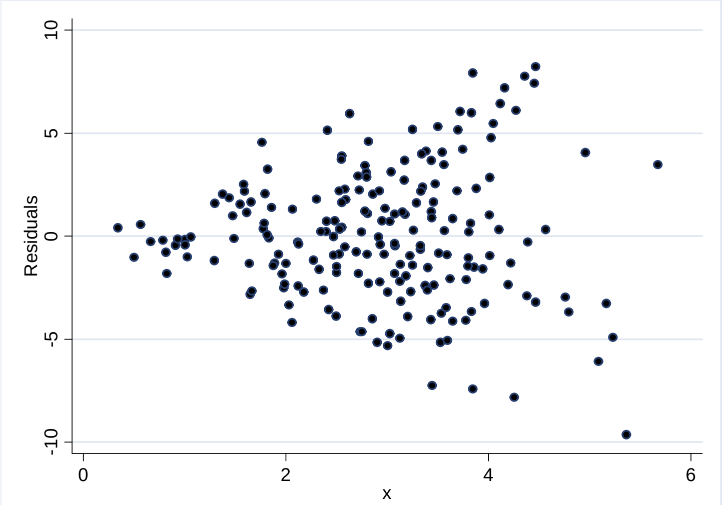
- Consider the following true regression model with heteroskedastic errors:

$$Y_i = 1 + 2X_i + u_i, \text{ where } u_i \sim N\left(0, X_i^2\right)$$



How Does the Heteroskedasticity Look?

- Alternatively, we can graph the residuals $\widehat{u}_i$ with X_i. Is there a constant spread across values of X?



Testing for Heteroskedasticity: Breusch-Pagan Test

- Assume that heteroskedasticity is of the linear form of independent variables:

$$\sigma_i^2 = \delta_0 + \delta_1 X_{i1} + \cdots + \delta_k X_{ik}.$$

Testing for Heteroskedasticity: Breusch-Pagan Test

- Assume that heteroskedasticity is of the linear form of independent variables:
$$\sigma_i^2 = \delta_0 + \delta_1 X_{i1} + \cdots + \delta_k X_{ik}.$$
- The hypotheses are $H_0 : Var\left(u_i | X_i\right) = \sigma^2$ and H_1 : not H_0. The null can be written
$$H_0 : \delta_1 = \cdots = \delta_k = 0.$$

Testing for Heteroskedasticity: Breusch-Pagan Test

- Assume that heteroskedasticity is of the linear form of independent variables:

$$\sigma_i^2 = \delta_0 + \delta_1 X_{i1} + \cdots + \delta_k X_{ik}.$$

- The hypotheses are $H_0 : Var\left(u_i | X_i\right) = \sigma^2$ and H_1 : not H_0. The null can be written

$$H_0 : \delta_1 = \cdots = \delta_k = 0.$$

- Since we never know the actual errors in the population model, we use their estimates, $\widehat{u}_i$, which is the OLS residual:

$$\widehat{u}_i^2 = \delta_0 + \delta_1 X_{i1} + \cdots + \delta_k X_{ik} + v_i.$$

Testing for Heteroskedasticity: Breusch-Pagan Test

- Assume that heteroskedasticity is of the linear form of independent variables:
$$\sigma_i^2 = \delta_0 + \delta_1 X_{i1} + \cdots + \delta_k X_{ik}.$$
- The hypotheses are $H_0 : Var\left(u_i | X_i\right) = \sigma^2$ and H_1 : not H_0. The null can be written
$$H_0 : \delta_1 = \cdots = \delta_k = 0.$$
- Since we never know the actual errors in the population model, we use their estimates, $\widehat{u}_i$, which is the OLS residual:
$$\widehat{u}_i^2 = \delta_0 + \delta_1 X_{i1} + \cdots + \delta_k X_{ik} + v_i.$$
- A form of the Breusch-Pagan test is constructed as
$$\text{BP test: } N \times R^2_{\widehat{u}^2} \sim^a \mathcal{X}_k^2.$$

Testing for Heteroskedasticity: White Test

- The White test is explicitly intended to test for forms of heteroskedasticity: the relation of u^2 with all independent variables (X_i), the squares of th independent variables $\left(X_i^2\right)$, and all the cross products $(X_i X_j$ for $i \neq j)$.

Testing for Heteroskedasticity: White Test

- The White test is explicitly intended to test for forms of heteroskedasticity: the relation of u^2 with all independent variables (X_i), the squares of th independent variables $\left(X_i^2\right)$, and all the cross products $(X_i X_j$ for $i \neq j)$.
- Just as we did in the Breusch-Pagan test, we regress $\widehat{u}_i$ on all the above variables and compute the $R^2_{\widehat{u}^2}$ and construct the statistic of same form.

Testing for Heteroskedasticity: White Test

- The White test is explicitly intended to test for forms of heteroskedasticity: the relation of u^2 with all independent variables (X_i), the squares of th independent variables $\left(X_i^2\right)$, and all the cross products $(X_i X_j \text{ for } i \neq j)$.
- Just as we did in the Breusch-Pagan test, we regress $\widehat{u}_i$ on all the above variables and compute the $R^2_{\widehat{u}^2}$ and construct the statistic of same form.
- The abundance of independent variables is a weakness in the pure form of the White test.

Heteroskedasticity-Robust Standard Errors

- Consider the simple regression model, $Y_i = \beta_0 + \beta_1 X_i + u_i$, and allow heteroskedasticity.
- Then, note that the variance of $\widehat{\beta}_1$ is

$$Var\left(\widehat{\beta}_1|X\right) = \frac{\sum_{i=1}^{N}\left(X_i - \overline{X}\right)^2 \sigma_i^2}{\left\{\sum_{i=1}^{N}\left(X_i - \overline{X}\right)^2\right\}^2}.$$

- White (1980) suggested the following:
  - Get the OLS residual $\widehat{u}_i$.
  - Get a valid estimator of $Var\left(\widehat{\beta}_1|X\right)$:

$$\widehat{Var\left(\widehat{\beta}_1|X\right)} = \frac{\sum_{i=1}^{N}\left(X_i - \overline{X}\right)^2 \widehat{u}_i^2}{\left\{\sum_{i=1}^{N}\left(X_i - \overline{X}\right)^2\right\}^2}.$$

Generalized Least Squares Estimation

- If we correctly specify the form of the variance, then there exists a more efficient estimator (Generalized Least Squares, GLS) than OLS.
- Suppose the true model is:

$$Y_i = \beta_0 + \beta_1 X_i + u_i, \;\; Var\left(u_i | X\right) = \sigma_i^2.$$

- Suppose we know exactly the form of heteroskedasticity. Then we divide each term of the equation by σ_i:

$$\begin{aligned} Y_i/\sigma_i &= \beta_0/\sigma_i + \beta_1 X_i/\sigma_i + u_i/\sigma_i \\ Y_i^* &= \beta_0^* + \beta_1 X_i^* + u_i^*, \;\; Var\left(u_i^* | X\right) = 1 \end{aligned}$$

- Perform the OLS regression of Y_i^* on X_i^*:

$$\widehat{\beta}_1^{GLS} = \frac{\sum_{i=1}^N \left(X_i^* - \overline{X}^*\right)\left(Y_i^* - \overline{Y}^*\right)}{\sum_{i=1}^N \left(X_i^* - \overline{X}^*\right)^2}$$

- In GLS, less weight is given the observations with a higher error variance. Obviously, GLS is unbiased and, indeed, is BLUE.

Feasible GLS

- The problem is we usually do not know the form of variance, σ_i.

Feasible GLS

- The problem is we usually do not know the form of variance, σ_i.
- Instead of σ_i, we can use $\widehat{\sigma}_i$ in the GLS estimation, called the **Feasible GLS (FGLS) estimator**.

Feasible GLS

- The problem is we usually do not know the form of variance, σ_i.
- Instead of σ_i, we can use $\widehat{\sigma}_i$ in the GLS estimation, called the **Feasible GLS (FGLS) estimator**.
  - Run the OLS regression to get the residuals, $\widehat{u}_i$.

Feasible GLS

- The problem is we usually do not know the form of variance, σ_i.
- Instead of σ_i, we can use $\widehat{\sigma}_i$ in the GLS estimation, called the **Feasible GLS (FGLS) estimator**.
  - Run the OLS regression to get the residuals, $\widehat{u}_i$.
  - Model the relation of errors with independent variables:

$$\sigma_i^2 = f(X_i)$$

Estimate $\widehat{\sigma}_i$ using the following OLS regression:

$$\widehat{u}_i^2 = f(X_i) + v_i$$

Feasible GLS

- The problem is we usually do not know the form of variance, σ_i.
- Instead of σ_i, we can use $\widehat{\sigma}_i$ in the GLS estimation, called the **Feasible GLS (FGLS) estimator**.
  - Run the OLS regression to get the residuals, $\widehat{u}_i$.
  - Model the relation of errors with independent variables:

    $$\sigma_i^2 = f\left(X_i\right)$$

    Estimate $\widehat{\sigma}_i$ using the following OLS regression:

    $$\widehat{u}_i^2 = f\left(X_i\right) + v_i$$

  - The feasible GLS estimator is

    $$\widehat{\beta}_1^{FGLS} = \frac{\sum_{i=1}^{N}\left(X_i^* - \overline{X}^*\right)\left(Y_i^* - \overline{Y}^*\right)}{\sum_{i=1}^{N}\left(X_i^* - \overline{X}^*\right)^2},$$

    where $X_i^* = X_i/\widehat{\sigma}_i$ and $Y_i^* = Y_i/\widehat{\sigma}_i$.

Autocorrelation

- The error terms are said to be autocorrelated if and only if

$$Cov\left(u_i, u_j\right) \neq 0, \quad \text{for } i \neq j.$$

- (**Time Series Data**) The error term at one date can be correlated with the error terms in the previous periods:
  - Autoregressive process of order $k = 1, 2, ...,$

$$AR\left(k\right): \;\; u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \cdots + \rho_k u_{t-k} + v_t.$$

  - Moving average process of order $k = 1, 2, ...,$

$$MA\left(k\right): \;\; u_t = v_t + \lambda_1 v_{t-1} + \cdots + \lambda_k v_{t-k}.$$

- (**Cross-section Data**) The error terms may be correlated with each other in terms of socio and geographical distance such as the distance between towns and neighborhood effects.

Consequences of Autocorrelation

- Assuming all other assumptions remain to hold, under the condition of autocorrelation,
  - the OLS estimator is still unbiased.
  - the OLS is not BLUE any more.
- The usual OLS standard errors and test statitics are no longer valid.
- We can find an autocorrelation-robust estimator of the variance after we perform the OLS regression.
- Alternatively, we can devise an *efficient* estimator by re-weighting the data appropriately to take into account of autocorrelation.

Autocorrelation-robust Standard Errors

- In order to correct the unknown form of autocorrelation in the error terms, Newey and West (1987) suggested

$$
\begin{aligned}
\widehat{Var\left(\widehat{\beta}_1 | X\right)} &= \frac{1}{\left\{\sum_{t=1}^{N}\left(X_t - \overline{X}\right)^2\right\}^2} \times \left\{\sum_{t=1}^{N} \widehat{u}_t^2 \left(X_t - \overline{X}\right)^2 \right. \\
&\quad \left. + \sum_{l=1}^{L}\sum_{t=l+1}^{N} w_l \widehat{u}_t \widehat{u}_{t-l}\left(X_t - \overline{X}\right)\left(X_{t-l} - \overline{X}\right)\right\},
\end{aligned}
$$

where

$$
w_l = 1 - \frac{l}{L+1}.
$$

- The correlation between u_t and u_{t-l} is approximated with $\left(1 - \frac{l}{L+1}\right)\widehat{u}_t\widehat{u}_{t-1}$.
- The above standard error is also robust to arbitrary heteroskedasticity.

Testing for Autocorrelation: AR(1)

- We start to test the presence of AR(1) serial correlation: $u_t = \rho u_{t-1} + e_t$.
- Then, the nulll hypothesis that the errors are serially uncorrelated is

$$H_0 : \rho = 0.$$

- In order to test the null,
  - run the OLS regression of Y_t on $X_{t1}, ..., X_{tk}$ to obtain the OLS residuals, $\widehat{u}_t$
  - run the regression of $\widehat{u}_t$ on $\widehat{u}_{t-1}$ for all $t = 2, ..., N$ to estimate $\widehat{\rho}$
  - construct the t statistic.
- The t test is not valid if

$$Cov\left(u_{t-1}, X_{tj}\right) \neq 0, \text{ for some } j.$$

Testing for Autocorrelation: AR(1)

- Another test for AR(1) serial correlation is the Durbin-Watson test: based on the OLS residuals,

$$\begin{aligned} d &= \frac{\sum_{t=2}^{N}\left(\widehat{u}_t - \widehat{u}_{t-1}\right)^2}{\sum_{t=1}^{N}\widehat{u}_t^2} = 2\left(1 - r\right) + \frac{\widehat{u}_1^2 + \widehat{u}_N^2}{\sum_{t=1}^{N}\widehat{u}_t^2} \\ &\simeq 2\left(1 - \widehat{\rho}\right), \text{ where } r = \frac{\sum_{t=2}^{N}\widehat{u}_t\widehat{u}_{t-1}}{\sum_{t=1}^{N}\widehat{u}_t^2}. \end{aligned}$$

- The DW test works as follows with two critical values, d_L and d_U:
  - $d \in [d_U, 4 - d_U] \Longrightarrow$ not reject the null;
  - either $d \leq d_L$ or $d \geq 4 - d_L \Longrightarrow$ reject the null;
  - $d \in (d_L, d_U)$ or $d \in (4 - d_U, 4 - d_L) \Longrightarrow$ inconclusive
- The DW test is also not valid if

$$Cov\left(u_{t-1}, X_{tj}\right) \neq 0, \text{ for some } j.$$

  - An important case is the regression with a lagged dependent variable:

$$Y_t = \beta_0 + \beta_1 X_t + \beta_2 Y_{t-1} + u_t, \;\; u_t = \rho u_{t-1} + e_t$$

Testing for Autocorrelation: Breusch-Godfrey Test

- The Breusch-Godfrey(BG) test is more general and test for higher order serial correlation, $AR(q)$:

$$u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \cdots + \rho_q u_{t-q} + v_t.$$

- Then, the null hypothesis is

$$H_0 : \rho_1 = \cdots = \rho_q = 0.$$

- In order to construct the BG test,
    - run the OLS regression of Y_t on $X_{t1}, ..., X_{tk}$ to obtain the OLS residuals, $\widehat{u}_t$
    - regress $\widehat{u}_t$ on $X_{t1}, ..., X_{tk}, \widehat{u}_{t-1}, ..., \widehat{u}_{t-k}$
    - compute the F statistic
    - Alternatively, compute $(N-q) R^2_{\widehat{u}}$, which follows the chi-square with q degrees of freedom.

GLS under Autocorrelation

- Consider the following model:

$$Y_t = \beta_0 + \beta_1 X_t + u_t \, , \;\; u_t = \rho u_{t-1} + v_t$$

- Assuming we know this relation, we can rewrite

$$Y_t - \rho Y_{t-1} = \beta_0 \left(1 - \rho\right) + \beta_1 \left(X_t - \rho X_{t-1}\right) + v_t$$

  and run the OLS regression.
- If ρ is not known, then we can do the feasible GLS in the following way:
  - run the OLS with the original equation to get $\widehat{u}_t$
  - run the regression: $\widehat{u}_t = \rho \widehat{u}_{t-1} + w_t$, to get $\widehat{\rho}$
  - transform the model using $\widehat{\rho}$ and run OLS.