Heteroskedasticity worksheet

Ryan T. Godwin

Heteroskedasticity

1. Mathematically define heteroskedasticity and homoskedasticity.

   Homoskedasticity is when:

   $$\text{Var}\left[\epsilon_i \mid X_i\right] = \sigma^2, \forall i$$

   and heteroskedasticity is when:

   $$\text{Var}\left[\epsilon_i \mid X_i\right] = \sigma_i^2$$

2. Which assumption does heteroskedasticity violate?

   A.4.

3. What does $V[\boldsymbol{\epsilon}]$ look like under heteroskedasticity?

   $V[\boldsymbol{\epsilon}] = \sigma^2\Omega = \Sigma$, where:

   $$\sigma^2\Omega = \sigma^2 \begin{bmatrix} \omega_{11} & 0 & \cdots & 0 \\ 0 & \omega_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \omega_{nn} \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n^2 \end{bmatrix}$$

4. Is the LS estimator unbiased under heteroskedasticity?

   The LS estimator is unbiased and consistent under heteroskedasticity. The proofs do not require the use of A.4.

5. What is $V[\boldsymbol{b}]$ under heteroskedasticity?

   $$V(\boldsymbol{b}) = \left[\left(X'X\right)^{-1} X'\sigma^2\Omega X \left(X'X\right)^{-1}\right]$$

   This means that the usual variance-covariance estimator $s^2(X'X)^{-1}$ based on homoskedasticity, is inconsistent under heteroskedasticity.

6. Is LS efficient under heteroskedasticity?

   LS is inefficient. A better estimator would make use of the fact that some observations have lower variance than others (and are more valuable).

7. What is the main problem here?

   The main problem is the inconsistency of standard errors (which come from the $s^2(X'X)^{-1}$ matrix). t-statistics, p-values, confidence intervals will all be incorrect. Hypothesis testing is invalid.

White's het. consistent covariance matrix (robust standard errors)

8. If we knew $V[\boldsymbol{\epsilon}]$, what would the estimator for $V[\boldsymbol{b}]$ be?

   The estimator would be:

$$\left[\left(X'X\right)^{-1} X'\sigma^2\Omega X \left(X'X\right)^{-1}\right]$$

9. Why can' we estimate $V[\boldsymbol{\epsilon}] = \Sigma$ directly? What can we estimate instead?

   Σ is $(n \times n)$, so we can't estimate it (there are too many parameters). However, $X'\sigma^2\Omega X$ is $(k \times k)$ and doesn't grow as n grows, so we can estimate it.

10. What is Σ replaced with in White's estimator?

    White's estimator essentially replaces each diagonal element of Σ with a squared residual.

11. What part of the standard regression output changes when a heteroskedastic consistent covariance matrix estimator (such as White's) is used?

    The estimated β do not change. The standard errors, t-statistics, and p-values, are all adjusted.

White's test for heteroskedasticity

12. What is the null in White's test for heteroskedasticity?

$$H_0 : \sigma_i^2 = \sigma^2$$

13. What is the basic idea behind the test?

    To see if there is any way to explain variation in the e_i^2. If there is, reject the null.

14. Explain how to conduct the test.

    (1) Estimate the model by LS, and get the residuals. (2) Using LS again, regress the e_i^2 values on each of the x's in the original model; their squared values; all of the cross-products of the regressors; and an intercept. (3) Get the nR^2 from step (2). Reject H_0 if the p-value for the nR^2 statistic from the chi-square distribution is small.