CHAPTER 8

What If the Disturbances Have Nonzero Expectations or Different Variances?

8.0 What We Need to Know When We Finish This Chapter

The ε_i's must have the same expected value for our regression to make any sense. However, we can't tell if the ε_i's have a constant expected value that is different from zero, and it doesn't make any substantive difference. If the disturbances have different variances, ordinary least squares (OLS) estimates are still unbiased. However, they're no longer best linear unbiased (BLU). In addition, the true variances of b and a are different from those given by the OLS variance formulas. In order to conduct inference, either we can estimate their true variances, or we may be able to get BLU estimators by *transforming* the data so that the transformed disturbances share the same variance. Here are the essentials.

1. **Section 8.2:** If $E(\varepsilon_i)$ equals some constant other than zero, b is still an unbiased estimator of β and a is still an unbiased estimator of the fixed component of the deterministic part of y_i.
2. **Section 8.2:** An *identification problem* arises when we don't have a suitable estimator for a parameter whose value we would like to estimate.
3. **Section 8.2:** A *normalization* is a value that we assign to a parameter when we don't have any way of estimating it and when assigning a value doesn't have any substantive implications.
4. **Section 8.2:** If $E(\varepsilon_i)$ equals some constant other than zero, it wouldn't really matter and we couldn't identify this constant. Therefore, we always normalize it to zero.
5. **Section 8.3:** If $E(\varepsilon_i)$ is different for each observation, then the observations don't come from the same population. In this case, b and a don't estimate anything useful.
6. **Section 8.4:** When the disturbances don't all have the same variance, it's called *heteroscedasticity*.
7. **Section 8.4:** *The OLS estimators b and a remain unbiased for* β *and* α *regardless of what we assume for* $V(\varepsilon_i)$.
8. **Section 8.5:** The OLS estimators b and a are not BLU and their true variances are probably not estimated accurately by the OLS variance formulas.
9. **Section 8.6:** An *auxiliary regression* does not attempt to estimate a population relationship in an observed sample. It provides supplemental information that helps us interpret regressions that do.
10. **Section 8.6:** The *White test* identifies whether heteroscedasticity is bad enough to distort OLS variance calculations.
11. **Equation (8.15), section 8.7:** The White heteroscedasticity-consistent variance estimator for b is

$$\mathrm{V_W}(b) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2 e_i^2}{\left(\sum (x_i - \bar{x})^2\right)^2}.$$

It and the corresponding variance estimator for a are consistent even if heteroscedasticity is present.

12. **Equation (8.19), section 8.8:** *Weighted least squares* (WLS) provides BLU estimators for β and α if the different disturbance variances are known or can be estimated:

$$\frac{y_i}{s_i} = a_{\mathrm{WLS}}\frac{1}{s_i} + b_{\mathrm{WLS}}\frac{x_i}{s_i} + e_i.$$

13. **Section 8.10:** Heteroscedasticity can take many forms. Regardless, the White heteroscedasticity-consistent variance estimator provides trustworthy estimates of the standard deviations of OLS estimates. In contrast, WLS estimates require procedures designed specifically for each heteroscedastic form.