POLS 6481. Research Design and Quantitative Methods II
Lecture 12. Circumventing Heteroskedasticity
Readings: Wooldridge, *Introductory Econometrics 5e,* 8.2 + 8.4

Outline

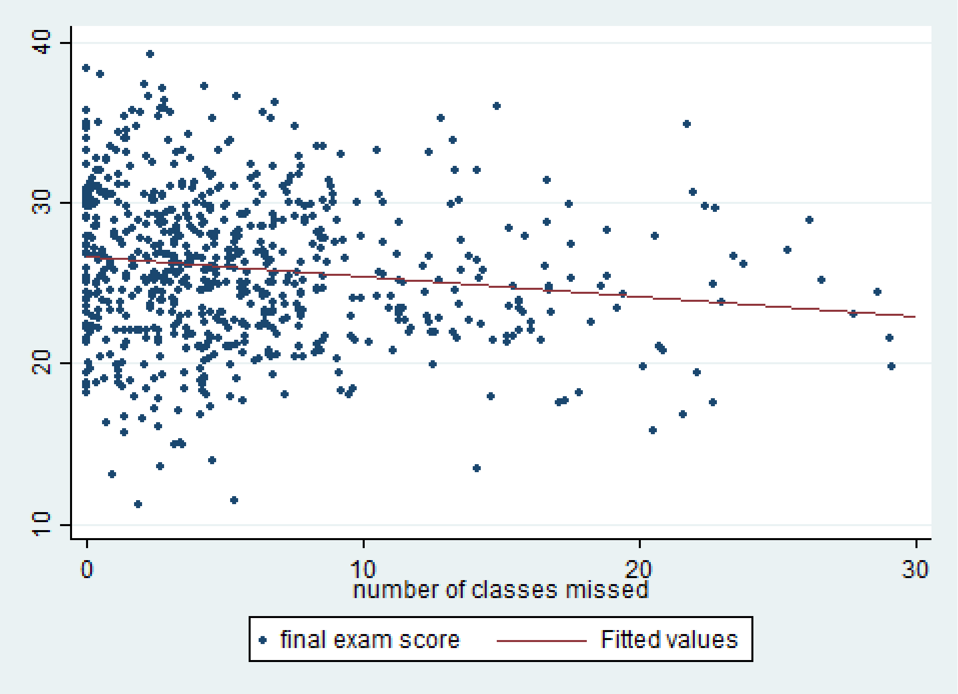
1. Example: *Lecture Attendance and Final Exam Scores*
2. Modeling Heteroskedasticity with Weighted Least Squares
3. Circumventing Heteroskedasticity with Robust Standard Errors
4. Extra Example: *Income and Demand for Cigarettes*

POLS 6481. Research Design and Quantitative Methods II
Lecture 12. Circumventing Heteroskedasticity
Readings: Wooldridge, *Introductory Econometrics 5e,* 8.2 + 8.4

1. Example: *Lecture Attendance and Final Exam Scores*

R scripts: Lecture 12 attendance.R & white-test.R
Dataset is attend-new.dta

Simple regression predicts final exam score from (non-) attendance in lecture



POLS 6481. Research Design and Quantitative Methods II
Lecture 12. Circumventing Heteroskedasticity
Readings: Wooldridge, *Introductory Econometrics 5e,* 8.2 + 8.4

1. Example: *Lecture Attendance and Final Exam Scores*

R scripts: Lecture 12 attendance.R & white-test.R
Dataset is attend-new.dta

Simple regression predicts final exam score from attendance in lecture

Multiple regression predicts final exam score from attendance, ACT score, GPA

- Examine bivariate relationship between attendance and final exam score
- Estimate regression model, predicting final exam score from attendance
- Carry out Breusch-Pagan test and/or White tests
- Compare ordinary standard errors to robust standard errors; compare *t* statistics
- Estimate feasible generalized least squares model (see p. 286–8) two ways :
  - regression with weights
  - regression with all variables (including constant) rescaled

POLS 6481. Research Design and Quantitative Methods II
Lecture 12. Circumventing Heteroskedasticity
Readings: Wooldridge, *Introductory Econometrics 5e,* 8.2 + 8.4

2. Modeling Heteroskedasticity with Weighted Least Squares

A. Weighted Least Squares

Suppose we know some function $h(\boldsymbol{x})$ that acts as a multiplier:

$$var(\, u \mid \boldsymbol{x}\,) = \sigma^2 h(\boldsymbol{x}) \qquad [8.21]$$

For a random sample from the population, we would write:

$$\sigma_i^2 = var(\, u_i \mid \boldsymbol{x}_i) = \sigma^2 h(\boldsymbol{x}_i) = \sigma^2 h_i$$

Take the original equation with heteroskedastic errors:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_k x_{ki} + u_i \qquad [8.24]$$

and transform it into one with homoskedastic errors:

$$\frac{y_i}{\sqrt{h_i}} = \beta_0 \frac{1}{\sqrt{h_i}} + \beta_1 \frac{x_{1i}}{\sqrt{h_i}} + \beta_2 \frac{x_{2i}}{\sqrt{h_i}} + \ldots + \beta_k \frac{x_{ki}}{\sqrt{h_i}} + \frac{u_i}{\sqrt{h_i}} \qquad [8.25]$$

This works because $\frac{u_i}{\sqrt{h_i}}$ has a zero expected value (conditional on $\boldsymbol{x}_i$) and variance equal to:

$$\mathrm{E}\left(\left(\frac{u_i}{\sqrt{h_i}}\right)^2\right) = \mathrm{E}\left(\frac{u_i^2}{h_i}\right) = \frac{E(u_i^2)}{h_i} = \frac{\sigma^2 h_i}{h_i} = \sigma^2$$

This is called **weighted** least squares because the coefficients minimized the **weighted** sum of squared residuals, where each squared residual is weighted by h_i^{-1} (p. 282–283)

2. Modeling Heteroskedasticity with Weighted Least Squares

B. Feasible Generalized Least Squares

Suppose we do not know the function $h(\boldsymbol{x})$ that acts as a multiplier; then estimate it:

$$var(\, u \mid \boldsymbol{x}\,) = \sigma^2 \cdot exp(\delta_0 + \delta_1 x_1 + \delta_2 x_2 + \ldots + \delta_k x_k)$$

$$\rightarrow \quad u^2 = \sigma^2 \cdot exp(\delta_0 + \delta_1 x_1 + \delta_2 x_2 + \ldots + \delta_k x_k) \cdot v$$

$$\rightarrow \quad log(u^2) = \mathrm{a}_0 + \delta_1 x_1 + \delta_2 x_2 + \ldots + \delta_k x_k + e$$

1. regress y on $x_1, x_2, \ldots, x_k$ for any positive integer k
2. generate fitted values ($\hat{y}_i$) and residuals ($\hat{u}_i$)
3. regress $log(\hat{u}^2)$ on $x_1, x_2, \ldots, x_k$ (or on $\hat{y}$ and $\hat{y}^2$) *
4. generate fitted values ($\hat{g}_i$)
5. exponentiate the fitted values: $\hat{h}_i = exp(\hat{g}_i)$
6. estimate the equation shown below, using weights $\hat{h}_i$ in place of h_i :

$$\frac{y_i}{\sqrt{\hat{h}_i}} = \beta_0 \frac{1}{\sqrt{\hat{h}_i}} + \beta_1 \frac{x_{1i}}{\sqrt{\hat{h}_i}} + \beta_2 \frac{x_{2i}}{\sqrt{\hat{h}_i}} + \ldots + \beta_k \frac{x_{ki}}{\sqrt{\hat{h}_i}} + \frac{u_i}{\sqrt{\hat{h}_i}}$$

Note that these estimates are not *unbiased* but they are *consistent* (asymptotically unbiased)

* An F test of the joint significance of the δ terms is called the Park test (p. 288)

POLS 6481. Research Design and Quantitative Methods II
Lecture 12. Circumventing Heteroskedasticity
Readings: Wooldridge, *Introductory Econometrics 5e,* 8.2 + 8.4

2. Modeling Heteroskedasticity with Weighted Least Squares

On the previous two slides, the models you are estimating regress a transformed dependent variable ($\frac{y_i}{\sqrt{\hat{h}_i}}$) on a transformed constant ($\frac{1}{\sqrt{\hat{h}_i}}$) and transformed independent variables ($\frac{x_{ji}}{\sqrt{\hat{h}_i}}$).

An alternative approach is simply to create weights equal to $\frac{1}{\sqrt{\hat{h}_i}}$ for each i.

You have some choices about which method to adopt; here are two opinions:

(1) tell *R* to weigh the observations is easier than transforming all the variables

(2) use robust standard errors, which entails re-estimating the standard errors after carrying out the regression; this is easier but allegedly is less efficient than FGLS

Readings: Wooldridge, *Introductory Econometrics 5e,* 8.2 + 8.4

3. Circumventing Heteroskedasticity with Robust Standard Errors

A. Compare variances

For k regressors, the estimated error variance is:

$$\hat{\sigma}^2 = \frac{1}{n-k-1}\sum \widehat{u_i}^2 \qquad [3.56]$$

In simple regression, the sampling variance of the slope (β) is:

$$\widehat{var}(\hat{\beta}) = \frac{\sum(x_i-\bar{x})^2\widehat{\sigma_i}^2}{(\sum(x_i-\bar{x})^2)^2} = \frac{\sum(x_i-\bar{x})^2\widehat{\sigma_i}^2}{{SST_x}^2} \qquad [8.2]$$

If we assume homoskedasticity, then $\widehat{\sigma_i} = \hat{\sigma}$ for all i, and then [8.2] is equivalent to:

$$\widehat{var}(\hat{\beta}) = \frac{\hat{\sigma}^2}{\sum(x_i-\bar{x})^2} = \frac{\hat{\sigma}^2}{SST_x} \qquad [2.57]$$

If we do not assume homoskedasticity, then the sampling variance of the slope (β) is:

$$\widehat{var}(\hat{\beta}) = \frac{\sum(x_i-\bar{x})^2 \cdot \widehat{u_i}^2}{{SST_x}^2} \qquad [8.3]$$

3. Circumventing Heteroskedasticity with Robust Standard Errors

A. Compare variances

In multiple regression, the sampling variance of a coefficient $\widehat{\beta}_j$ is :

$$\widehat{var}(\widehat{\beta}_j) = \frac{\widehat{\sigma}^2}{\sum\left(x_{ji}-\overline{x_j}\right)^2 \cdot \left[1-R_j^2\right]} \qquad [3.54]$$

where R_j^2 is the R^2 from regressing x_j on all $x_{\sim j}$ and where $\sum\left(x_{ji} - \overline{x_j}\right)^2 = SST_{x_j}$

If we do not assume homoskedasticity, then a coefficient's sampling variance becomes:

$$\widehat{var}(\widehat{\beta}_j) = \frac{\sum\left(\widehat{r_{ij}}\right)^2 \cdot \widehat{u_i}^2}{SSR_j^2} \qquad [8.4]$$

where $\widehat{r_{ij}}$ is the i^{th} residual from regressing x_j on all $x_{\sim j}$ and where SSR_j^2 is the *residual sum of squares* from this regression. Recall however that $SSR = (1–R^2) \cdot SST$…

… so if R_j^2 is the R^2 from this regression, then $SSR_j^2 = (1–R_j^2) \cdot SST_{x_j}$, and therefore:

$$\widehat{var}(\widehat{\beta}_j) = \frac{\sum\left(\widehat{r_{ij}}\right)^2 \cdot \widehat{u_i}^2}{\left(SST_{x_j} \cdot \left[1-R_j^2\right]\right)^2} = \frac{\sum\left(\widehat{r_{ij}}\right)^2 \cdot \widehat{u_i}^2}{SST_{x_j}{}^2 \cdot \left[1-R_j^2\right]^2}$$

POLS 6481. Research Design and Quantitative Methods II
Lecture 12. Circumventing Heteroskedasticity
Readings: Wooldridge, *Introductory Econometrics 5e,* 8.2 + 8.4

3. Circumventing Heteroskedasticity with Robust Standard Errors

B. Heteroskedasticity–robust t tests

Recall that the standard error of a coefficient is the root of the estimated variance, i.e.,

$$se(\widehat{\beta}_j) = \sqrt{\widehat{var}(\widehat{\beta}_j)}$$

Easy version of the t statistic: $t = \dfrac{\widehat{\beta_j} - a_j}{robust\ se(\widehat{\beta_j})}$ for some null hypothesized value a_j.

Harder version of the t statistic: multiply $\widehat{var}(\widehat{\beta}_j)$ by $\frac{n}{n-k-1}$ before taking the root.

(Wooldridge also discusses heteroskedasticity–robust F and LM tests in 8.2)

POLS 6481. Research Design and Quantitative Methods II
Lecture 12. Circumventing Heteroskedasticity
Readings: Wooldridge, *Introductory Econometrics 5e,* 8.2 + 8.4

4. Extra Example: *Income and Demand for Cigarettes*

R scripts: Lecture 12 cigarettes.R and white-test.R
Dataset is SMOKE.DTA

Simple regression (illustration only) predicts cigarette consumption from *log*(income)

Multiple regression predicts cigarette consumption from *log*(income), *log*(cigarette price), years of education, and dummy variable for state smoking restrictions in restaurants

- Estimate baseline model (drop *age* and age^2 from Example 8.7, p. 288–9)
- Carry out Breusch-Pagan test and White test (Wooldridge's special version in [8.20])
  note: pay attention to F statistic not to *t* statistics because of multicollinearity
- Compare ordinary standard errors to robust standard errors; compare *t* statistics
- Estimate feasible generalized least squares model (see p. 286–8) two ways :
  - regression with weights
  - regression with all variables (including constant) rescaled