Econometrics-Damodar N. Gujarati / Chapter 11

Furkan Zengin

20 08 2021

Heteroscedasticity

$$Homoscedasticity \qquad (11.1.1)$$

$$E(u_i^2) = \sigma^2$$

```
library(pacman)
p_load(
  broom, latex2exp, ggplot2, ggthemes, viridis, extrafont,
  dplyr,
  magrittr, knitr, parallel
)
# Define pink color
red_pink <- "#e64173"
grey_light <- "grey70"
grey_mid <- "grey50"
grey_dark <- "grey20"
# Dark slate grey: #314f4f
# Notes directory
dir_slides <- "~/Dropbox/UO/Teaching/EC421W19/LectureNotes/02Review/"
# Knitr options
opts_chunk$set(
  comment = "#>",
  fig.align = "center",
  fig.height = 7,
  fig.width = 10.5,
  warning = F,
  message = F
)
# A blank theme for ggplot
theme_empty <- theme_bw() + theme(
  line = element_blank(),
  rect = element_blank(),
  strip.text = element_blank(),
  axis.text = element_blank(),
  plot.title = element_blank(),
  axis.title = element_blank(),
  plot.margin = structure(c(0, 0, -0.5, -1), unit = "lines", valid.unit = 3L, class = "unit"),
  legend.position = "none"
)
theme_simple <- theme_bw() + theme(
  line = element_blank(),
  panel.grid = element_blank(),
  rect = element_blank(),
  strip.text = element_blank(),
  axis.text.x = element_text(size = 18, family = "STIXGeneral"),
  axis.text.y = element_blank(),
  axis.ticks = element_blank(),
  plot.title = element_blank(),
  axis.title = element_blank(),
  # plot.margin = structure(c(0, 0, -1, -1), unit = "lines", valid.unit = 3L, class = "unit"),
  legend.position = "none"
)
theme_axes_math <- theme_void() + theme(
  text = element_text(family = "MathJax_Math"),
  axis.title = element_text(size = 22),
  axis.title.x = element_text(hjust = .95, margin = margin(0.15, 0, 0, 0, unit = "lines")),
  axis.title.y = element_text(vjust = .95, margin = margin(0, 0.15, 0, 0, unit = "lines")),
  axis.line = element_line(
    color = "grey70",
```

```
    size = 0.25,
    arrow = arrow(angle = 30, length = unit(0.15, "inches")
  )),
  plot.margin = structure(c(1, 0, 1, 0), unit = "lines", valid.unit = 3L, class = "unit"),
  legend.position = "none"
)
theme_axes_serif <- theme_void() + theme(
  text = element_text(family = "MathJax_Main"),
  axis.title = element_text(size = 22),
  axis.title.x = element_text(hjust = .95, margin = margin(0.15, 0, 0, 0, unit = "lines")),
  axis.title.y = element_text(vjust = .95, margin = margin(0, 0.15, 0, 0, unit = "lines")),
  axis.line = element_line(
    color = "grey70",
    size = 0.25,
    arrow = arrow(angle = 30, length = unit(0.15, "inches")
  )),
  plot.margin = structure(c(1, 0, 1, 0), unit = "lines", valid.unit = 3L, class = "unit"),
  legend.position = "none"
)
theme_axes <- theme_void() + theme(
  text = element_text(family = "Fira Sans Book"),
  axis.title = element_text(size = 18),
  axis.title.x = element_text(hjust = .95, margin = margin(0.15, 0, 0, 0, unit = "lines")),
  axis.title.y = element_text(vjust = .95, margin = margin(0, 0.15, 0, 0, unit = "lines")),
  axis.line = element_line(
    color = grey_light,
    size = 0.25,
    arrow = arrow(angle = 30, length = unit(0.15, "inches")
  )),
  plot.margin = structure(c(1, 0, 1, 0), unit = "lines", valid.unit = 3L, class = "unit"),
  legend.position = "none"
)
```

This code Belongs to Edward Rubin:https://github.com/edrubin/EC421S20

```
set.seed(12345)
ggplot(data = tibble(
  x = runif(1e3, -3, 3),
  e = rnorm(1e3, 0, sd = 4 + 1.5 * x)
), aes(x = x, y = e)) +
geom_point(color = "darkslategrey", size = 2.75, alpha = 0.5) +
labs(x = "x", y = "u") +
theme_axes_math
```

```
set.seed(12345)
ggplot(data = tibble(
  g = sample(c(F,T), 1e3, replace = T),
  x = runif(1e3, -3, 3),
  e = rnorm(1e3, 0, sd = 0.5 + 2 * g)
), aes(x = x, y = e, color = g, shape = g, alpha = g)) +
geom_point(size = 2.75) +
scale_color_manual(values = c("darkslategrey", red_pink)) +
scale_shape_manual(values = c(16, 1)) +
scale_alpha_manual(values = c(0.5, 0.8)) +
labs(x = "x", y = "u") +
theme_axes_math
```

```
set.seed(12345)
# Data
gq_df <- tibble(
  x = runif(1e3, -3, 3),
  e = rnorm(1e3, 0, sd = 4 + 1.5 * x),
  y = 1 + 3 * x + e
)
# Quantiles
gq_x <- quantile(gq_df$x, probs = c(3/8, 5/8))
# Regressions
sse1 <- lm(y ~ x, data = gq_df %>% filter(x < gq_x[1])) %>%
  residuals() %>% raise_to_power(2) %>% sum()
sse2 <- lm(y ~ x, data = gq_df %>% filter(x > gq_x[2])) %>%
  residuals() %>% raise_to_power(2) %>% sum()
ggplot(data = gq_df, aes(x = x, y = e)) +
geom_point(color = "darkslategrey", size = 2.75, alpha = 0.5) +
labs(x = "x", y = "u") +
theme_axes_math
```

```
ggplot(data = gq_df, aes(
  x = x, y = e,
  color = cut(x, c(-Inf, gq_x, Inf)),
  alpha = cut(x, c(-Inf, gq_x, Inf)),
  shape = cut(x, c(-Inf, gq_x, Inf))
)) +
geom_vline(
  xintercept = gq_x,
  color = grey_mid,
  size = 0.25
) +
geom_point(size = 2.75) +
labs(x = "x", y = "u") +
scale_color_manual(values = c("darkslategrey", grey_mid, red_pink)) +
scale_shape_manual(values = c(19, 1, 19)) +
scale_alpha_manual(values = c(0.5, 0.8, 0.6)) +
theme_axes_math
```

OLS Estimation in the Presence of Heteroscedasticity

$$\hat{\beta}_2 = \frac{\sum_i (x_i y_i)}{\sum_i (x_i^2)} \tag{11.2.1}$$

$$\text{Var}(\hat{\beta}_2) = \frac{\sum_i \left(x_i^2 \sigma_i^2\right)}{\sum_i (x_i^2)^2} \tag{11.2.2}$$

The Method of Generalized Least Squares

$$Y_i = \beta_1 X_{0i} + \beta_2 X_i + u_i \quad \text{and} \quad \frac{Y_i}{\sigma_i} = \beta_1 \frac{X_{0i}}{\sigma_i} + \beta_2 \frac{X_i}{\sigma_i} + \frac{u_i}{\sigma_i} \tag{11.3.3 and 11.3.4}$$

Detection of Heteroscedasticity

Goldfeld–Quandt Test

```
options(scipen = 999)

library(gujarati)

library(ggplot2)

library(lmtest)
fix(Table11_3)

attach(Table11_3)

MODEL1= lm(Table11_3$Y~Table11_3$X)
summary(MODEL1)

gqtest(MODEL1, order.by = ~Table11_3$X,fraction = 13)
```

Breusch–Pagan–Godfrey Test

$$p_i = \alpha_1 + \alpha_2 Z\_2i + \alpha_3 Z\_3i + \cdots + \alpha_m Z_{mi} + v_i \tag{11.5.15}$$

$$\text{LM} = n \times R_e^2$$

```
MODEL1= lm(Table11_3$Y~Table11_3$X)
summary(MODEL1)
bptest(MODEL1)

# For the White test

bptest(MODEL1,~fitted(MODEL1)+I(fitted(MODEL1)^2))
```

Remedial Measures

The Method of Weighted Least Squares

```
# Since the data in R are not compatible with the book,
# We can use wooldridge data.

library(wooldridge)

View(k401ksubs)

attach(k401ksubs)

#Using OLs

MODEL2=lm(nettfa~inc+I((age-25)^2)+male+e401k,subset=(fsize==1))

summary(MODEL2)

#Using WLS

MODEL3=lm(nettfa~inc+I((age-25)^2)+male+e401k,subset=(fsize==1),weights =
1/inc)

summary(MODEL3)

library(stargazer)

stargazer(MODEL2,MODEL3,type="text",column.labels = c("OLS","WLS"))

plot(inc,nettfa)

abline(MODEL2,lwd=3,lty=1,col="red")

abline(MODEL3,lwd=3,lty=3,col="green")

legend("topleft",c("OLS","WLS"),lty =c(1,3),bty = "n")


#When the weight is not known we can create by the process:

attach(smoke)

#Standard OLS

MODEL4=lm(cigs~lincome+lcigpric+educ+age+agesq+restaurn)

summary(MODEL4)

#Now we need residuals of previous regression

weight1=lm(log(residuals(MODEL4)^2)~lincome+lcigpric+educ+age+agesq+restaurn,
data = smoke)

MODEL5=lm(cigs~lincome+lcigpric+educ+age+agesq+restaurn,weights =
1/exp(fitted(weight1)),data = smoke)

summary(MODEL5)
```

```
stargazer(MODEL4,MODEL5,type="text",column.labels = c("OLS","WLS"))

#If we consider weight that we have chosen is wrong
#Then we can look at the Robust WLS.

library(lmtest)
library(sandwich)

robust3=vcovHC(MODEL7,type = "HC1")

MODEL3_Robust=coeftest(MODEL3,robust3)

stargazer(MODEL3,MODEL3_Robust,type = "text",column.labels = c("Non-Robust WLS","Robust WLS"))
```

Empirical Exercises

11.15

```


fix(Table11_7)


MODEL7 = lm(Table11_7$MPG ~ Table11_7$HP +Table11_7$WT+ Table11_7$SP)
summary(MODEL7)

bptest(MODEL7)

MODEL8 = lm(log(Table11_7$MPG) ~ log(Table11_7$HP) +log(Table11_7$WT)+ log(Table11_7$SP))
summary(MODEL8)

bptest(MODEL8)

#When we take log values, it becomes homoskedastic

stargazer(list(MODEL7,MODEL8),type = "text")
```