

## Aufgabenblatt 2

### Lösungen

---

## Überblick

*Heteroskedastie*

*Paneldaten*

## Aufgaben

1. Bitte schätzen Sie das Modell

$$\text{cigs} = \beta_0 + \beta_1 \log(\text{income}) + \beta_2 \log(\text{cigpric}) + \beta_3 \text{educ} + \beta_4 \text{age} + \beta_5 \text{age}^2 + \beta_6 \text{restaurn} + u \quad (1)$$

mit dem Datensatz `wooldridge::smoke` und interpretieren Sie die Koeffizienten.

```
R> ols_smoke <- lm(
+   formula = cigs ~ lincome + lcigpric + educ + age + I(age^2) + restaurn,
+   data = wooldridge::smoke
+ )
R> summary(ols_smoke)

#
# Call:
# lm(formula = cigs ~ lincome + lcigpric + educ + age + I(age^2) +
#     restaurn, data = wooldridge::smoke)
#
# Residuals:
#      Min       1Q   Median       3Q      Max
# -15.819   -9.381   -5.975    7.922   70.221
#
# Coefficients:
#              Estimate Std. Error t value Pr(>|t|)
# (Intercept) -3.639841   24.078660  -0.151  0.87988
# lincome      0.880268    0.727783   1.210  0.22682
# lcigpric    -0.750859    5.773343  -0.130  0.89655
# educ        -0.501498    0.167077  -3.002  0.00277 **
# age          0.770694    0.160122   4.813 1.78e-06 ***
# I(age^2)    -0.009023    0.001743  -5.176 2.86e-07 ***
# restaurn    -2.825085    1.111794  -2.541  0.01124 *
# ---
# Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#
```

```
# Residual standard error: 13.4 on 800 degrees of freedom
# Multiple R-squared:  0.05274, Adjusted R-squared:  0.04563
# F-statistic: 7.423 on 6 and 800 DF,  p-value: 9.499e-08
```

- $\beta_0$ : durchschnittliche Anzahl gerauchter Zigaretten pro Tag für alle Regressoren gleich Null (ergibt hier keinen Sinn und ist nicht signifikant)
- $\beta_1$ : für eine Erhöhung des Jahreseinkommens um 1% impliziert das Modell eine Erhöhung der Anzahl durchschnittlich gerauchter Zigaretten pro Tag um approximativ 0.01 Zigaretten, ceteris paribus (nicht signifikant).
- $\beta_2$ : für eine Erhöhung des Packungspreises um 1% impliziert das Modell eine Verminderung der Anzahl der durchschnittlich gerauchten Zigaretten pro Tag um 0.008 Zigaretten, ceteris paribus (nicht signifikant).
- $\beta_3$ : für ein weiteres Jahr Schulbildung impliziert das Modell eine Verminderung der Anzahl der durchschnittlich gerauchten Zigaretten pro Tag um eine halbe Zigarette, ceteris paribus.
- $\beta_4$  und  $\beta_5$ : der partielle Effekt ist  $\beta_4 + 2\beta_5 \text{age}$ , hält man also alle andere Variablen fix, so impliziert das Modell, dass ein um ein Jahr höheres Alter zu einer Veränderung der durchschnittlich gerauchten Zigaretten pro Tag in Höhe von approximativ  $0.8 - 0.02 \times \text{age}$  führt.
- $\beta_6$ : für eine Regulierung des Rauchens in Restaurants impliziert das Modell eine Verminderung der durchschnittlich gerauchten Zigaretten pro Tag in Höhe von knapp drei Zigaretten unter sonst gleichen Bedingungen.

## 2. Liegt in Modell (1) Heteroskedastie vor?

```
R> ols_smoke_white <- lm(
+   formula = ols_smoke$residuals^2 ~ ols_smoke$fitted + I(ols_smoke$fitted^2)
+ )
R> summary(ols_smoke_white)

#
# Call:
# lm(formula = ols_smoke$residuals^2 ~ ols_smoke$fitted + I(ols_smoke$fitted^2))
#
# Residuals:
#      Min       1Q   Median       3Q      Max
# -328.0  -108.6   -94.3   -62.3  4732.4
#
# Coefficients:
#              Estimate Std. Error t value Pr(>|t|)
# (Intercept)      14.0534     47.7985   0.294    0.769
# ols_smoke$fitted      14.0534     11.5674   1.215    0.225
# I(ols_smoke$fitted^2)   0.4920      0.7556   0.651    0.515
#
# Residual standard error: 363.7 on 804 degrees of freedom
```

```
# Multiple R-squared:  0.03293, Adjusted R-squared:  0.03052
# F-statistic: 13.69 on 2 and 804 DF,  p-value: 1.427e-06
```

Der White-Test sagt ja!

3. Verwenden Sie statt OLS den FGLS-Schätzer um (1) zu schätzen.

Vorgehensweise der FGLS-Schätzung wird in der Vorlesung beschrieben:

1. schätze wie gewohnt mit OLS und speichere die Residuen  $u$
2. berechne  $\log(u)$
3. schätze das Modell  $\log(u) = \gamma_0 + \gamma'X + v$
4. berechne die Gewichte  $1/\exp(\hat{\gamma}_0 + \hat{\gamma}'X)$
5. berechne das Modell aus Schritt 1 erneut mit den Gewichten (einfach die Gewichte an das `weights` Argument der `lm()` Funktion übergeben)

```
R> u <- residuals(ols_smoke)
R> model_helper <- lm(
+   formula = I(log(u^2)) ~ lincome + lcigpric + educ + age + I(age^2) + restaurn,
+   data = wooldridge::smoke
+ )
R> weights <- 1 / exp(model.matrix(model_helper) %*% coefficients(model_helper))
R> fgls_smoke <- lm(
+   formula = cigs ~ lincome + lcigpric + educ + age + I(age^2) + restaurn,
+   data = wooldridge::smoke,
+   weights = weights
+ )
R> summary(fgls_smoke)

#
# Call:
# lm(formula = cigs ~ lincome + lcigpric + educ + age + I(age^2) +
#     restaurn, data = wooldridge::smoke, weights = weights)
#
# Weighted Residuals:
#      Min       1Q   Median       3Q      Max
# -1.9036 -0.9532 -0.8099  0.8415  9.8556
#
# Coefficients:
#              Estimate Std. Error t value Pr(>|t|)
# (Intercept)  5.6354618  17.8031385   0.317  0.751673
# lincome      1.2952399   0.4370118   2.964  0.003128 **
# lcigpric     -2.9403123   4.4601445  -0.659  0.509930
# educ         -0.4634464   0.1201587  -3.857  0.000124 ***
# age          0.4819479   0.0968082   4.978  7.86e-07 ***
```

310214 Praktische Übung zu “Einführung in die Mikroökonomie”

```
# I(age^2)    -0.0056272  0.0009395  -5.990 3.17e-09 ***
# restaurn    -3.4610641  0.7955050  -4.351 1.53e-05 ***
# ---
# Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#
# Residual standard error: 1.579 on 800 degrees of freedom
# Multiple R-squared:  0.1134, Adjusted R-squared:  0.1068
# F-statistic: 17.06 on 6 and 800 DF,  p-value: < 2.2e-16
```