

Econ 444 Elementary Econometrics

Fall 2007

SUMMARIES AND SAMPLE PROBLEMS for

CHAPTERS 6, 7, 8, 10

CHAPTER 6

KEY TERMS AND CONCEPTS

Omitted variables, Omitted variable bias, Irrelevant variables

FORMULAS AND RESULTS

The adjusted R^2 , denoted by the symbol \bar{R}^2 is defined by

$$\bar{R}^2 = 1 - (1 - R^2) \frac{N - 1}{N - K - 1} \quad (1)$$

CHAPTER 7

KEY TERMS AND CONCEPTS Double-log functional form, Semilog functional form, Dummy variables

CHAPTER 8

KEY TERMS AND CONCEPTS Perfect multicollinearity, Imperfect multicollinearity

CHAPTER 10

KEY TERMS AND CONCEPTS Homoskedasticity, Heteroskedasticity, Weighted Least Squares (WLS)

FORMULAS AND RESULTS

Consider the following regression:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \epsilon_i \quad (2)$$

Assume that heteroskedasticity is caused by a proportionality factor Z_i where $Z_i > 0$:

$$Var(\epsilon_i) = \sigma_i^2 = \sigma^2 Z_i^2. \quad (3)$$

where σ^2 is a constant. Consider the following transformation of the model

$$Y_i^* = \beta_0 \frac{1}{Z_i} + \beta_1 X_i^* + u_i \quad (4)$$

where

$$\begin{aligned} Y_i^* &= \frac{Y_i}{Z_i} \\ X_i^* &= \frac{X_i}{Z_i} \\ u_i &= \frac{\epsilon_i}{Z_i} \end{aligned} \tag{5}$$

The OLS estimators of β_0 and β_1 obtained by applying the OLS to (4) are called *weighted least squares estimators* (WLS).

SAMPLE PROBLEMS

1. Consider a regression

$$\ln Y_i = \beta_0 + \beta_1 \ln X_{1i} + \beta_2 \ln X_{2i} + \epsilon_i \tag{6}$$

where Y is the demand for food consumption, X_1 is the price of food, and X_2 is income. Imagine that you use a cross-section data set of 18 households. Answer the following questions.

(a) To test the hypothesis $H_0 : \beta_1 = 0$ with a one-tailed test, which alternative hypothesis is appropriate? $H_A : \beta_1 > 0$ or $H_A : \beta_1 < 0$? Explain.

(b) For the one-tailed test in part (a), answer whether or not you reject the null hypothesis for each of the following combinations of hypothetical values of $\hat{\beta}_1$ and $SE(\hat{\beta}_1)$ at the 5% level.

(b-1) $\hat{\beta}_1 = -1.5$ and $SE(\hat{\beta}_1) = 0.5$

(b-2) $\hat{\beta}_1 = -1.5$ and $SE(\hat{\beta}_1) = 0.9$

(b-3) $\hat{\beta}_1 = -1.5$ and $SE(\hat{\beta}_1) = 1.2$

(b-4) $\hat{\beta}_1 = 1.5$ and $SE(\hat{\beta}_1) = 0.5$

(b-5) $\hat{\beta}_1 = 1.5$ and $SE(\hat{\beta}_1) = 0.9$

(b-6) $\hat{\beta}_1 = 1.5$ and $SE(\hat{\beta}_1) = 1.2$

(c) To test the hypothesis of unit elasticity for food against the alternative hypothesis of elastic demand with a one-tailed test, which alternative hypothesis is appropriate, $H_A : \beta_1 > -1$ or $H_A : \beta_1 < -1$? Explain.

(d) For the one-tailed test in part (c), answer whether or not you reject the null hypothesis for each of the the following combinations of hypothetical values of $\hat{\beta}_1$ and $SE(\hat{\beta}_1)$ at the 5% level.

- (d-1) $\hat{\beta}_1 = -2.1$ and $SE(\hat{\beta}_1) = 0.5$
- (d-2) $\hat{\beta}_1 = -2.1$ and $SE(\hat{\beta}_1) = 0.9$
- (d-3) $\hat{\beta}_1 = -2.1$ and $SE(\hat{\beta}_1) = 1.2$
- (d-4) $\hat{\beta}_1 = -1.4$ and $SE(\hat{\beta}_1) = 0.15$
- (d-5) $\hat{\beta}_1 = -0.5$ and $SE(\hat{\beta}_1) = 0.15$

(e) To test the hypothesis $H_0 : \beta_2 = 0$ with a one-tailed test, which alternative hypothesis is appropriate, $H_A : \beta_2 > 0$ or $H_A : \beta_2 < 0$? Explain.

(f) For the one-tailed test in part (e), answer whether or not you reject the null hypothesis for each of the the following combinations of hypothetical values of $\hat{\beta}_2$ and $SE(\hat{\beta}_2)$ at the 5% level.

- (f-1) $\hat{\beta}_2 = 1.5$ and $SE(\hat{\beta}_2) = 0.5$
- (f-2) $\hat{\beta}_2 = 1.5$ and $SE(\hat{\beta}_2) = 0.9$
- (f-3) $\hat{\beta}_2 = 1.5$ and $SE(\hat{\beta}_2) = 1.2$
- (f-4) $\hat{\beta}_2 = 0.5$ and $SE(\hat{\beta}_2) = 0.15$
- (f-5) $\hat{\beta}_2 = -1.5$ and $SE(\hat{\beta}_2) = 0.15$

2. Consider the following two regressions on annual time series data for demand for food, the price of food, and the total expenditures from 1941 to 1962.

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \epsilon_i \quad (7)$$

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 D_i + \epsilon_i \quad (8)$$

where Y is the log of the budget share for food (the expenditures on food over the total consumption expenditures), X_2 is the log of the price of food, and X_3 is the log of the total real expenditures, and D_i is the dummy variable for World War II. The dummy variable takes the value of one for the period of 1941 to 1945. Suppose that you obtain the following results as computer output:

Regression (7): without the dummy
 Regression Output:
 Constant 0.231490
 R Squared 0.758182
 X Coefficient(s) -0.03278 -0.01946
 Std Err of Coef. 0.015325 0.008287

Regression (8): with the dummy

Regression Output:

Constant 0.263148

R Squared 0.761226

X Coefficient(s) -0.03122 -0.02325 -0.37761

Std Err of Coef. 0.015869 0.012315 0.558459

- (a) Report the two regression results with our class format.
- (b) Interpret $\hat{\beta}_1$ of Regression (7).
- (c) Test the hypothesis $\beta_1 = 0$ at the 5% significance level, using an appropriate one-tailed test for Regression (7). Which alternative hypothesis do you use for the one-tailed test?
- (d) Engel's law states that the budget share for food decreases as consumers become richer. Test the hypothesis $B_2 = 0$ against the alternative hypothesis that Engel's law holds, using Regression (7) at the 5% significance level. Which test do you use, two-tailed or one-tailed test? Explain.
- (e) Test the hypothesis that all regression coefficients are zero except for the intercept for Regression (8) at the 1 % significance level.
- (f) What is the food budget share predicted by the regression (8) when X_1 is 15, X_2 is 60 during World War II . What is the value of the share predicted by Regression (8) for these X_1 and X_2 values after World War II?
- (g) Which regression do you prefer, the regression (7) or the regression (8)? Explain.

3. What tends to happen to standard errors and t -value when a regression has multicollinearity?

4. In the regression, $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$, suppose that the error variance has the following structure:

$$E(\epsilon_i^2) = \sigma^2 X_i^4. \quad (9)$$

How would you transform the data to achieve homoskedastic error variance? How would you estimate the transformed model? List the various steps.