



# B.E. International Program

Faculty of Economics, Thammasat University



Semester: 1/2015

EE 425 Econometrics 1

Homework #5 (Heteroscedasticity and Autocorrelation)

1. In a survey of some 9966 economists in 1964 the following data were obtained:

Age, years	Median salary, \$
20-24	7,800
25-29	8,400
30-34	9,700
35-39	11,500
40-44	13,000
45-49	14,800
50-54	15,000
55-59	15,000
60-64	15,000
65-69	14,500
70+	12,000

Source: "The Structure of Economists' Employment and Salaries," Committee on the National Science Foundation Report on the Economics Profession, *American Economic Review*, vol. 55, no. 4, December 1965, p.36.

- (a) Estimate a linear regression model explaining median salary in relation to age. Evaluate your result.

Note: For the purpose of regression assume that the median salaries refer to the midpoint of the age interval. Thus, \$7800 refers to age 22 years, and so on. For the last age interval, assume that the maximum age is 75 years.

- (b) Assuming that the variance of the disturbance term is proportional to the square of age, transform the data so as to make the resulting disturbance term homoscedastic. Estimate the regression and evaluate the result.
- (c) Repeat (b) assuming that the variance is proportional to age. Which of the transformations do you prefer? Provide your reasons.

2. In a regression of average wages (W) on the number of employees (N) for a random sample of 30 firms, the following regression results were obtained:

$$\bar{W}_i = 7.5 + 0.009N_i \quad (1)$$

$$t = (3.0) (16.10) \quad R^2 = 0.90$$

$$\bar{W}_i / N_i = 0.008 + 7.8(1/N_i) \quad (2)$$

$$t = (14.43) (76.58) \quad R^2 = 0.99$$

- (a) How do you interpret the two regressions?  
 (b) What is the author assuming in going from Eq. (1) to (2)? Was he worried about heteroscedasticity? How do you know?  
 (c) Can you relate the slopes and intercepts of the two models?  
 (d) Can you compare the  $R^2$  values of the two models? Why or why not?
3. You are given the accompanying data:

Y, Personal consumption expenditure (billions of 1958 dollars)	X, time	$\hat{Y}$ , estimated $Y^*$	$\hat{u}$ , residuals
281.4	1 (= 1956)	261.4208	19.9791
288.1	2	276.6026	11.4973
290.0	3	291.7844	-1.7844
307.3	4	306.9661	0.3338
316.1	5	322.1479	-6.0479
322.5	6	337.3297	-14.8297
338.4	7	352.5115	-14.1115
353.3	8	367.6933	-14.3933
373.7	9	382.8751	-9.1751
397.7	10	398.0569	-0.3569
418.1	11	413.2386	4.8613
430.0	12	428.4206	1.6795
452.7	13	443.6022	9.0977
469.1	14	458.7840	10.3159
476.9	15 (= 1970)	473.9658	2.9341

\* Obtained from the regression model:  $Y_t = \beta_0 + \beta_1 X_t + u_t$ .

- (a) Calculate the Durbin-Watson d statistic.  
 (b) Is there positive serial correlation in the disturbances? Provide your reasons.

4. In studying the movement in the production workers' share in the value of output (i.e., labor's share), the following models were considered by Gujarati:<sup>\*</sup>

$$\text{Model A: } Y_t = \beta_0 + \beta_1 t + u_t$$

$$\text{Model B: } Y_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + u_t$$

where  $Y$  = labor's share and  $t$  = time. Based on annual data for 1949-1964 the following results were obtained for the primary metal industry:

$$\begin{aligned} \text{Model A: } \hat{Y}_t &= 0.4529 - 0.0041t \\ &\quad (-3.9608) \\ R^2 &= 0.5284 \quad d = 0.8252 \end{aligned}$$

$$\begin{aligned} \text{Model B: } \hat{Y}_t &= 0.4786 - 0.0127t + 0.0005t^2 \\ &\quad (-3.2724) \quad (2.7777) \\ R^2 &= 0.6629 \quad d = 1.82 \end{aligned}$$

where the figures in the parentheses are  $t$  ratios.

- (a) Is there serial correlation in model A? In model B?  
 (b) What do you think accounts for the serial correlation?

---

\* Damodar Gujarati, "Labor's Share in Manufacturing Industries," *Industrial and Labor Relations Review*, vol. 23, no.1, October 1969, pp.65-75.