Chapter 7

1. \*What is the dummy variable about? Give an example using dummy in the regression and explain the model.
2. \*There is a give model below. Explain the case when d=1 and d=0 in the graph. What is the additional value in the model when d=1?
3. \*\*We have the following regression model:

(235.11) (0.018) (5.86) (11.21) (0.134) (34.33)

n=706; R-squared=0.123;

t-critical value of 5% confidence level for this model is 1.984 and the values under each coefficient are their standard errors.

The variable sleep is total minutes per week spent sleeping at night, totwrk is total weekly minutes spent working, educ and age are measured in years, and male is a gender dummy. All other factors being equal, is there evidence that men sleep more than women? How strong is the evidence?

1. \*\*Having the following model:

(0.22) (0.009) (0.0059) (0.006) (0.010) (0.013)

n= 1 388; R-squared=0.0472

t-critical value of 5% confidence level for this model is 1.984 and the values under each coefficient are their standard errors.

How much more is a white child predicted to weigh than a nonwhite child, holding the other factors in the first equation fixed? Is the difference statistically significant?

1. \*\*Given model is:

(6.29) (3.83) (0.53) (4.29) (12.71) (18.15)

n=4137; R-squared=0.0858

The variable sat is the combined SAT score, hsize is size of the student’s high school graduating class, in hundreds, female is a gender dummy variable and black is a race dummy variable equal to one for blacks and zero otherwise.

Holding hsize fixed, what is the estimated difference in SAT score between nonblack females and nonblack males?

1. \*\* Consider the model

n=6763; R-squared=0.202

Using this equation, find the value of totcoll such that the predicted values of log(wage) are the same for men and women.

Chapter 8

1. \* What is heteroskedasticity? Why is it important to have it, although there is unbiased and consistent OLS estimates without this assumption?
2. \*\*What are the consequences of heteroskedasticity? Explain it clearly.
3. \*\*\*Prove that , when there is homoscedasticity.
4. \*\*\*Derive the equation of  when there is heteroskedasticy in the regression.
5. \*\*Explain how to test the heteroskedasticity? What are the well-known tests for that to use and their differences?
6. \*\*\*Consider a linear model to explain monthly beer consumption:

Write the transformed equation that has a homoskedastic error term and explain how you came up with that equation.

1. \*\*The variable smokes is a binary variable equal to one if a person smokes, and zero otherwise. We estimate a linear probability model for smokes:

At what point does another year of age reduce the probability of smoking? Explain the way you find it.

1. \*\*\*There are different ways to combine features of the Breusch-Pagan and White tests for heteroskedasticity. One possibility not covered in the text is to run the regression

where the are the OLS residuals and the are the OLS fitted values. Then, we would test joint significance of Xi1, Xi2, …, Xik and . (Of course, we always include an intercept in this regression.)

Explain why the R-squared from the regression above will always be at least as large as the R-squareds for the BP regression and the special case of the White test.

Chapter 9

1. \*What tests are used in order to detect general functional form misspecification? Explain the conduction of the test on the example.
2. \*What is the difference between nested and nonnested models? What tests are used to detect the right functional forms of these models? Explain it clearly.
3. \*How to test against nonnested models? What approaches there are? Explain them.
4. \*Consider that the model is misspecified due to the unavailable data of an independent variable. What method is needed to be used to solve the misspecification? Explain the method and the way solved the problem clearly.
5. Let math10 denote the percentage of students at a Michigan high school receiving a passing \*score on a standardized math test. We are interested in estimating the effect of per student spending on math performance. A simple model is

where poverty is the percentage of students living in poverty. The variable lnchprg is the percentage of students eligible for the federally funded school lunch program. Why is this a sensible proxy variable for poverty?

Chapter 10

1. \*What is the difference between time-series and cross sectional analysis? Give a data samples for both and show the difference on their data as well.
2. \*\*State the Gauss-Markov assumptions for time series analysis and explain each of them briefly. Show the difference of those with that of Gauss-Markov assumptions of cross sectional analysis as well.
3. \*Assume that the dependent and independent variables are trending together over time. Why the relationship is not always causal? Identify the reasons and solutions for that.
4. \*\*When the three event indicators befile, affile, and afdec are dropped from the below equation, we obtain R-squared = 0.281 and R(adjusted)-squared= 0 .264. Are the event indicators jointly significant at the 10% level, considering F-critical value is about 2.13?

n=132; R-squared = 0.305; R(adjusted)-squared= 0 .271.

Chapter 11

1. \*What are stationary and non-stationary processes about? Explain the difference and each element in it.
2. \*\*Explain the weakly dependent time series analysis in terms of covariance stationary process. Why it is important to have weak dependence in time series analysis?