

## Chapter 2

1. Let  $kids$  denote the number of children ever born to a woman, and let  $educ$  denote years of education for the woman. A simple model relating fertility to years of education is

$$kids = \beta_0 + \beta_1 educ + u,$$

where  $u$  is the unobserved error.

- i. What kinds of factors are contained in  $u$ ? Are these likely to be correlated with level of education?
  - ii. Will a simple regression analysis uncover the ceteris paribus effect of education on fertility? Explain.
- (i)  $u$  can be various factors that effect number of children ever bord to woman such as age, income, etc. Moreover,  $u$  is corrolated with level of education. For example, people who have high income are likely to have high education as well as have more children than people who have low income because they have enouge fund to take care their children.
- (ii) This simple regression analysis is uncover the ceteris paribus because not only the education can explain the number of children ever bord to woman but age and income can also explain it. [ $u$  is corrolated with level of education.]

4. The data set BWGHT contains data on births to women in the United States. Two variables of interest are the dependent variable, infant birth weight in ounces (*bwght*), and an explanatory variable, average number of cigarettes the mother smoked per day during pregnancy (*cigs*). The following simple regression was estimated using data on  $n = 1,388$  births:

$$\widehat{bwght} = 119.77 - 0.514 \text{ cigs}$$

- i. What is the predicted birth weight when *cigs* = 0? What about when *cigs* = 20 (one pack per day)? Comment on the difference.
  - ii. Does this simple regression necessarily capture a causal relationship between the child's birth weight and the mother's smoking habits? Explain.
  - iii. To predict a birth weight of 125 ounces, what would *cigs* have to be? Comment.
  - iv. The proportion of women in the sample who do not smoke while pregnant is about .85. Does this help reconcile your finding from part (iii)?
- (i)  $\widehat{bwght} = 119.77 - 0.514 \text{ Cigs}$   
 $\text{cigs} = 0 : \widehat{bwght} = 119.77 - 0.514(0) = 119.77$   
 $\text{cigs} = 20 : \widehat{bwght} = 119.77 - 0.514(20) = 109.49$   
 When women are not consume any cigarette, infant birth weight will be 119.77 ounce.  
 If women consume cigarettes one pack per day (*cigs* = 20), we expected that infant birth weight will decrease  $119.77 - 109.49 = 10.28$  ounces
- (ii) The amount of cigarettes that mother smoke is independent variables and weight of infant is dependent variable. So, this is a casual effect. But, there are other factors that need to be considered such as genetic and health.
- (iii)  $\widehat{bwght} = 119.77 - 0.514 \text{ Cigs}$   
 $125 = 119.77 - 0.514 \text{ cigs}$   
 $\text{cigs} = \frac{125 - 119.77}{-0.514} = -10.175$
- It is impossible the *cigs* in negative number (-10.175), the minimum *cigs* is equal to 0. So, the maximum possible infant birth weight is when *cigs* = 0 which equal to 119.77 ounce.
- (iv) In this situation, there is too less sample to predict. We need more sample of women who smoke to calculate the true sample to be more accurate. According to SLR2, the more sample is better. So, this is not as limited and inaccurate data.

1. Using the data in GPA2 on 4,137 college students, the following equation was estimated by OLS:

$$\widehat{colgpa} = 1.392 - .0135 hspc + .00148 sat$$

$$n = 4,137, R^2 = .273,$$

where  $colgpa$  is measured on a four-point scale,  $hspc$  is the percentile in the high school graduating class (defined so that, for example,  $hspc = 5$  means the top 5% of the class), and  $sat$  is the combined math and verbal scores on the student achievement test.

- Why does it make sense for the coefficient on  $hspc$  to be negative?
- What is the predicted college GPA when  $hspc = 20$  and  $sat = 1,050$ ?
- Suppose that two high school graduates, A and B, graduated in the same percentile from high school, but Student A's SAT score was 140 points higher (about one standard deviation in the sample). What is the predicted difference in college GPA for these two students? Is the difference large?
- Holding  $hspc$  fixed, what difference in SAT scores leads to a predicted  $colgpa$  difference of .50, or one-half of a grade point? Comment on your answer.

Then, student A have  $\widehat{colgpa} = 0.00148(140) = 0.2072$  higher than B  
 $\therefore$  This difference is quite large. compared to standard deviation

$$(iv) \widehat{colgpa} = 1.392 - 0.0135 \overline{hspc} + 0.00148 sat$$

$$\text{Difference of } \widehat{colgpa} = 0.00148 sat$$

$$0.5 = 0.00148 sat$$

$$sat = 337.8378$$

$\therefore$  sat must be different 337.8378 point in half of grade point.

2. The data in WAGE2 on working men was used to estimate the following equation:

$$\widehat{educ} = 10.36 - .094 sibs + .131 meduc + .210 feduc$$

$$n = 722, R^2 = .214,$$

where  $educ$  is years of schooling,  $sibs$  is number of siblings,  $meduc$  is mother's years of schooling, and  $feduc$  is father's years of schooling.

- Does  $sibs$  have the expected effect? Explain. Holding  $meduc$  and  $feduc$  fixed, by how much does  $sibs$  have to increase to reduce predicted years of education by one year? (A noninteger answer is acceptable here.)
- Discuss the interpretation of the coefficient on  $meduc$ .
- Suppose that Man A has no siblings, and his mother and father each have 12 years of education. Man B has no siblings, and his mother and father each have 16 years of education. What is the predicted difference in years of education between B and A?

$$(iii) \widehat{educ} = 10.36 - 0.094 sibs + 0.131 meduc + 0.210 feduc$$

$$\text{Man A } \widehat{educ} = 10.36 - 0.094(0) + 0.131(12) + 0.210(12)$$

$$= 10.36 + 1.572 + 2.52$$

$$= 14.452$$

$$\text{Man B } \widehat{educ} = 10.36 - 0.094(0) + 0.131(16) + 0.210(16)$$

$$= 10.36 + 2.096 + 3.36$$

$$= 15.816$$

$\therefore$  The difference of predicted years of education between B and A is  $15.816 - 14.452 = 1.364$  years

## Chapter 3

Thanyalak Samreeloy  
6104641284

(i) Because 'hspc' should be low due to the top of the class has low percentage. In addition, high 'colgpa' comes with top class student. So, coefficient is negative.

$$(ii) \widehat{colgpa} = 1.392 - 0.0135 hspc + 0.00148 sat$$

$$= 1.392 - 0.0135(20) + 0.00148(1050)$$

$$= 2.676$$

$\therefore$  Predicted college GPA is 2.676

$$(iii) \widehat{colgpa} = 1.392 - 0.0135 hspc + 0.00148 sat$$

Increasing in 1 point of sat will rise 'colgpa' up 0.00148