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 is the error terms of the observation, Let's say that age are contained in u and correlated with education. so, this equation cannot be OLS as violating SLR.4 $E(u_i|x) = 0$ which is not be true in this equation.

ii) Ceteris paribus is holding other things constant, only education cannot explain well about the fertility

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*:

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

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) $\widehat{bwght} = 119.77 - 0.514(0) = 119.77$

$\widehat{bwght} = 119.77 - 0.514(20) = 109.49$

the effect of smoking on additional 20 cigarettes with cause an estimated (119.77 − 109.49) decrease in the birth weight.

ii) As the amount of cigarettes is independent variable and the infant weight is dependent, it implies that there is a causal effect. However, there are other factors to be considered as well like genetics. Smoking habits can partially explain on baby's weight.

iii) $\hat{y} = 119.77 - 0.514 \text{ cigs}$

Cigs = -19.17s, which is not possible, the highest possible weight according to the model would be its β_0 which is 119.7s ounces.

Chapter 7

Q1) i) Yes, it does make sense if you are top of the class, the 'hsperc' will be low, and top class usually set high 'colgpa'. Hence, the coefficient is negative.

ii) 'colgpa' = 1.392 - 0.0135(20) + 0.0148(1053)

= 2.696

iii) 'colgpa' = 1.392 - 0.0135 hsperc + 0.00148 sat

→ thus, increasing in sat score for 1 mark will make the value of

$\widehat{\text{College}}$ change for 0.0019

$$\Rightarrow \frac{\partial \widehat{\text{college}}}{\partial \text{Sat}} = 0.0019$$

iv) Sat $\Rightarrow$ 37.839

Q2) i) yes, the higher the number of year, the lower the year schooling.

this is b/c the budget constraint,

$\Rightarrow 1 = -0.94 \text{ sibs}$

$\Rightarrow \text{sibs} = 1.064$ fo reduce yrs education

ii) If Mother's yrs of schooling ↑ by 1, the predicted years of school would increase by .13↑

iii) Man A predicted yrs of edu.

$= 10.76 - 0.94(0) + .171(2) + .210(12)$

$= 10.76 + 1.542 + 2.52 = 14.452$

Man B predicted yrs of edu.

$= 10.76 - 0.0944(0) + .13(16) + .210(16)$

$= 15.616$

∴ predicted different is 1.164