# Problem Set 6 Solutions

ECON 210 Econometrics A

Evan Zuofu Liao*

November 19, 2015

## Question 1 (W 3.3)

**(i)** Clearly $\beta_1$ is supposed to be negative given the assumption that adults trade off sleep for work.
**(ii)** The signs of $\beta_2$ and $\beta_3$ are ambiguous. The argument can go either way.
**(iii)** Note first that the unit of *totwrk* is minute, so we must convert hours into minutes: $\Delta$*totwrk*$= 5(60) = 300$. Then *sleep* is predicted to fall $\Delta$*sleep*$= .148(300) = 44.4$ minutes per week, which does not seem to be a substantial reduction.
**(iv)** If *educ* is measured in years, then an additional year of education predicts 11 minutes reduction in sleeping time per week, other things equal.
**(v)** These three variables explains about 11% of the total variation in sleeping time. There are certainly other factors playing a role here, for example, health. One's health would also generally be correlated with *totwrk*.

## Question 2 (W 3.4)

**(i)** This is obvious. A less prestige school has a larger value for *rank*, and has less starting salaries.
**(ii)** *LSAT* and *GPA* reflect the quality of the entering class, *libvol* and *cost* reflect the quality of school, so they are all supposed to be positively associated with salary.
**(iii)** The difference in salary is simply 24.8%
**(iv)** The coefficient can be interpreted as elasticity: a one percent increase in library volumes implies a 0.095% increase in predicted median starting salary, other things equal.

*Comments and questions to evanliao@uchicago.edu. This solution draws from answers provided by previous TAs.

**(v)** If a higher rank means a larger value for *rank*, then it is better to attend a school with a lower rank. The predicted difference in starting salary resulted from a rank difference of 20 is $(0.0033)20 = 6.6\%$

## Question 3 (W 3.9)

**(i)** We suspect that $\beta_1 < 0$ because more pollution is expected to lower housing prices and that $\beta_2 > 0$ because the number of rooms is a good proxy for the size of a house.
**(ii)** *rooms* and *nox* might be negatively correlated because higher *rooms* might suggest better neighborhoods that often have less pollution. If $\beta_2 > 0$ and $\text{Corr}\,(rooms, nox) < 0$, the simple regression estimator has a downward bias. In other words, it overestimates the importance of pollution.
**(iii)** A higher $\beta_1$ $(-0.718 > -1.043)$ matches with our prediction in part (ii). If we think the sample is representative, the true elasticity is closer to -0.718

## Question 4

**(a)** Looking simply at average does not allow us to conclude racial discrimination. It only tells us that black people as a sample had a higher average deny rate. However, this could be due to the fact that black people on average have lower income or prefer to buy homes with much higher value loans.

**(b)** Consider Model 1 below

$$\begin{aligned} Deny = \beta_0 + \beta_1(MP/Inc) + \beta_2(Loan/Val) + &\beta_3 Black + \beta_4(MP/Inc \times Black) \\ &+ \beta_5(Loan/Val \times Black) + \epsilon \end{aligned}$$

The null hypothesis is that race plays no role. In other words, we have $H_0 : \beta_3 = \beta_4 = \beta_5 = 0$ vs. $H_a : o.w.$. Follow the lecture, define $R$ to be a 3 by 6 restriction matrix. The null hypothesis becomes

$$\begin{aligned} H_0 :& R\beta = 0 \\ H_a :& R\beta \neq 0 \end{aligned}$$

$$\text{where } R = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \text{ and } \beta = [\beta_0\ \beta_1\ \beta_2\ \beta_3\ \beta_4\ \beta_5]'$$

Based on this hypothesis test, we can propose the following test statistic

$$T_n = (R\hat{\beta})'(R\widehat{\Sigma}_n R')^{-1}(R\hat{\beta})$$

Note that in order to compute $T_n$, we need to obtain the variance-covariance matrix $\widehat{\Sigma}_n$. However, there is not enough information provided in the table to let us get $\widehat{\Sigma}_n$, so we have to turn to other method to test the hypothesis.

Now look at Model 2 which excludes the interaction terms.

$$Deny = \alpha_0 + \alpha_1(MP/Inc) + \alpha_2(Loan/Val) + \alpha_3 Black + \varepsilon$$

Testing whether race has an impact on denial decision is equivalent to testing whether the partial derivative of deny with respect to race is zero. In Model 1, the partial derivative is $\beta_3 + \beta_4 MP/Inc + \beta_5 Loan/Val$, so we need to test if $\beta_3 = \beta_4 = \beta_5 = 0$. In model 2, this partial derivative is conveniently summarized by the parameter $\alpha_3$. So we can simply test if $\alpha_3 = 0$ using the standard test statistic for a single restriction

$$|T_n| = \frac{|\hat{\alpha}_3 - 0|}{\sqrt{\hat{\sigma}_3^2/n}} = \frac{|\hat{\alpha}_3|}{S.E.(\hat{\alpha}_3)} = \frac{0.63}{0.08} = 7.875$$

If the sample size is small, this test statistic follows a t-distribution with $(n-k-1)$ degrees of freedom where $k+1$ is the number of regressors. In this problem, since we have a relatively large sample size, so we can say that the test statistic follows a standard normal distribution.

**(c)** Reading off a standard normal distribution table, we find that the 5% critical value for a two-sided test is 1.96. Clearly we can reject the null hypothesis that race does not change evaluation of financial characteristics at the 5% level.

**(d)** In a linear regression model with binary dependent variable, coefficients can be interpreted as the change in probability that $Y = 1$ for a given change in $X$. In Model 2, the coefficient on black, 0.63, indicates that being black has a 63% higher probability of having a mortgage application denied, holding all other factors constant. We can get a similar estimate from Model 1 by computing $(0.71 + 0.68 * 0.33 - 0.39 * 0.74) = 64.6\%$.

**(e)** From part (c) we know that the discrimination effect is statistically significant. The estimate in part (d) suggests that there might also be an economically significant racial bias in morrgage decisions. However, such conclusion would be premature if there are some other factors playing a role in mortgage decision that are also correlated with the regressor *Black*. For instance, credit history might be an important variable that is omitted in the regression. If being black has a positive indirect effect on mortgage decision through credit history, then the coefficient in this model tend to overestimate the discrimination effect.

# Question 5

**(a)** In model 1 our coefficient, 20.19, captures how much sale prices increase due to an additional bedroom. In other words holding all other regressors fixed the sale price of a house would increase by $\$20,190$ if it hypothetically had an additional bedroom.

**(b)** The null hypothesis is $H_0 : \beta_3 = 0$ versus the alternative hypothesis that $H_a : \beta_3 \neq 0$. This is a simple t-test. Remember that the confidence interval is the non-rejection region, and suppose that the sample size is large enough, we have

$$|T_n| = \frac{|\hat{\beta}_3 - 0|}{\sqrt{\hat{\sigma}_3^2/n}} = \frac{|\hat{\beta}_3|}{S.E.} = \frac{(20.19)}{7.96} \leq Z_{0.975}$$
$$CI = \Big\{20.19 - 1.96 \times 7.96,\ 20.19 + 1.96 \times 7.96\Big\} = \Big\{4.5884,\ 35.7916\Big\}$$

where 1.96 is the 5% critical value for a two-sided test using a standard normal distribution. Clearly zero is not included in the confidence region, so we can reject the null hypothesis.

**(c)** Now the hypothesis becomes $H_0 : \beta_2 = \beta_5 = 0$ vs $H_a : o.w.$. Note again that the direct implementation of the F-test is not feasible given the information in Table 2. So we turn to the following equivalent F-test using Model 1 as the unrestricted model and Model 3 as the restricted model.

$$F_n = \frac{(SSR_r - SSR_{ur})/q}{SSR_{ur}/(n-k-1)} = \frac{(300720 - 222020)/2}{222020/(66-5-1)} \approx 10.63$$

The 5% critical value for an $F(2, 60)$ distribution is 3.15, so we can reject the null hypothesis that $\beta_2 = \beta_5 = 0$