**Econometrics I**

*Lab exercises-Chapter 2*

*Multiple Regression Exercises: Dummy Variables*

**2.1)** (Ex. 7.10 Wooldridge). Use the data in WAGE2.xls for this exercise, where *married*, *south*, *black* y *urban* are dummy variables.

(i) Estimate the model

$$\begin{aligned} log(wage) &= \beta_0 + \beta_1 educ + \beta_2 exper + \beta_3 tenure \\ &+ \beta_4 married + \beta_5 black + \beta_6 south + \beta_7 urban + u, \end{aligned}$$

and report the results in the usual form. Holding other factors fixed, what is the approximate difference in monthly salary between blacks and nonblacks? Is this difference statistically significant?

(ii) Add the variables $exper^2$ and $tenure^2$ to the equation and show that they are jointly insignificant at even the 20 % level.

(iii) Extend the original model to allow the return to education to depend on race and test whether the return to education does depend on race.

(iv) Again, start with the original model, but now allow wages to differ across four groups of people: married and black, married and nonblack, single and black, and single and nonblack. What is the estimated wage differential between married blacks and married nonblacks?

**2.2)** (Ex. 7.15 Wooldridge). Use the data in WAGE1.xls for this exercise. The model is the following:

$$\begin{aligned} log(wage) &= \beta_0 + \beta_1 female + \beta_2 educ + \beta_3 female \cdot educ \\ &+ \beta_4 exper + \beta_5 exper^2 + \beta_6 tenure + \beta_7 tenure^2 + u. \end{aligned}$$

(i) Estimate the gender differential when $educ = 12.5$. Compare this with the estimated differential when $educ = 0$.

(ii) Estimate the previous model but with $female \cdot (educ - 12.5)$ instead of $female \cdot educ$. How do you intepret the coefficient on female now?

(iii) Is the coefficient on $female$ in part (ii) statistically significant? Compare this with (i) and comment.

**2.3)** (Ex. 7.16 Wooldridge). Use the data in LOANAPP.xls for this exercise. The binary variable to be explained is *approve*, which is equal to one if a mortgage loan to an individual was approved. The key explanatory variable is *white*, a dummy variable equal to one if the applicant was white. The other applicants in the data set are *black* and *Hispanic*. To test for discrimination in the mortgage loan market, a linear probability model can be used:

$$approve = \beta_0 + \beta_1 white + other\ factors.$$

(i) If there is discrimination against minorities, and the appropriate factors have been controlled for, what is the sign of $\beta_1$?

(ii) Regress *approve* on *white* and report the results in the usual form. Interpret the coefficient on *white*. Is it statistically significant? Is it practically large?

(iii) As controls, add the variables *hrat*, *obrat*, *loanprc*, *unem*, *male*, *married*, *dep*, *sch*, *cosign*, *chist*, *pubrec*, *mortlat*1, *mortlat*2 and *vr*. What happens to the coefficient on *white*? Is there still evidence of discrimination against nonwhites?

(iv) Now allow the effect of race to interact with the variable measuring other obligations as a percent of income (*obrat*). Is the interaction term significant?

(v) Using the model from part (iv), what is the effect of being white on the probability of approval when $obrat = 32$, which is roughly the mean value in the sample? Obtain a 95 % confidence interval for this effect.