

# Homework

## Introduction to Econometrics

Fall 2005

### Instructions

Using the 526 observations on workers available from Wooldridge's site <http://fmwww.bc.edu/ec-p/data/wooldridge/WAGE1.dta>, answer each of the following. Remember, you can use the USE command in STATA, e.g.,

```
use http://fmwww.bc.edu/ec-p/data/wooldridge/WAGE1
```

#### Variables:

1. wage	average hourly earnings
2. educ	years of education
3. exper	years potential experience
4. tenure	years with current employer
5. nonwhite	=1 if nonwhite
6. female	=1 if female
7. married	=1 if married
8. numdep	number of dependents
9. smsa	=1 if live in SMSA
10. northcen	=1 if live in north central U.S
11. south	=1 if live in southern region
12. west	=1 if live in western region
13. construc	=1 if work in construc. indus.
14. ndurman	=1 if in nondur. manuf. indus.
15. trcompu	=1 if in trans, commun, pub ut
16. trade	=1 if in wholesale or retail
17. services	=1 if in services indus.
18. profserv	=1 if in prof. serv. indus.
19. profocc	=1 if in profess. occupation
20. clerocc	=1 if in clerical occupation
21. servocc	=1 if in service occupation
22. lwage	$\ln(\text{wage})$
23. expersq	$\text{exper}^2$
24. tenursq	$\text{tenure}^2$

## Questions

Consider the following linear model of log wages:

$$\ln(wage_i) = \beta_1 + \beta_2 educ_i + \beta_3 exper_i + \beta_4 tenure_i + u_i \quad (1)$$

where  $i = 1, 2, \dots, 526$ .

1. Estimate the model's parameters using least squares.
2. In tabular form, present the estimated coefficients, the usual least squares standard errors, and  $R^2$ .
3. Reestimate the model using heteroscedasticity robust standard errors. Add these results to your table and then compare them to the usual least squares results.
4. Using robust standard errors, test the hypothesis that education has a positive effect on wages at the 5% level. Are any other variables significantly positive at this level?
5. Using robust standard errors, form a 90% confidence interval for  $\beta_3$ . What does this indicate?
6. Obtain the predicted values from this regression and plot these, along with actual log wages, against education. (You'll need to sort the data using `educ` or the graph won't make sense). Could you have guessed what the graph would look like from the estimated coefficient? Why?