

# Homework

Econ 5243

November 29, 2004

## Problems

This assignment is due on the last day of class. Feel free to ask questions about this assignment in class.

Consider the following model of wages:

$$\ln(\text{wage}_i) = \beta_1 + \beta_2 \text{educ}_i + \beta_3 \text{exper}_i + \beta_4 \text{expersq}_i + \beta_5 \text{female}_i + u_i \quad (1)$$

where

`wage_i` = wages of the `i`th individual  
`educ_i` = years of schooling for individual `i`  
`exper_i` = years of experience for individual `i`  
`expersq_i` = years of experience squared for individual `i`  
`female_i` = 1 if female, 0 otherwise

also included in the data set is:  
`married_i` = 1 if married, 0 otherwise

1. Using the data

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

estimate the model using least squares.

- (a) Is the regression significant at the 5% level?
- (b) What is  $R^2$ ?

- (c) What does inclusion of *exper* and *expersq* accomplish? Do these have the anticipated signs?
- (d) What effect does an additional year of schooling have on wages?
- (e) Using equation (1), perform the RESET test at the 5% level. Is there evidence that the model is misspecified? Create a new variable, *educsq* as the square of *educ*. Augment the model by including this variable and repeat the RESET test. Does this improve things?
- (f) Jethro Bodean of Beverly Hill, California has a 6<sup>th</sup> grade education. What is the estimated impact of this on his log wage? Using the model that includes *educsq*, is there any incentive for Jethro to finish the 7<sup>th</sup> grade?
- (g) What is the estimated return to another year of schooling for someone who has finished high school (*educ* = 12<sup>th</sup>) and college (*educ* = 16<sup>th</sup>).
- (h) Is there evidence of gender discrimination in the model?
- (i) Suppose you believe that married women suffer more discrimination than single women. Reformulate the model to account for this and test the hypothesis at the 5% level.
- (j) Suppose you believe that single women earn the same wages as men, but that only married women are victims of discrimination. Reformulate the model in terms of this belief and test the hypothesis at the 5% level. Is this model ‘better’ than that in (f) above?
- (k) By now you have estimated 4 models i.e., (a), (e), (i), and (j). Put the results into a table and state which you prefer to base your conclusions on.

2. Consider the model

$$\ln(\text{wage}_i) = \beta_1 + \beta_2 \text{educ}_i + \beta_3 \text{educsq}_i + \beta_4 \text{exper}_i + \beta_5 \text{expersq}_i + \beta_6 \text{female}_i + \beta_7 \text{female}_i * \text{married}_i + u_i \quad (2)$$

- (a) Is there evidence that equation (2) is heteroscedastic? Be specific here by performing the appropriate hypothesis test.
- (b) Test the hypothesis that *tenure* and *exper* lead to heteroscedasticity in the errors of (2).
- (c) If you conclude that the model is heteroscedastic, are any of the results presented above valid? If so, which?
- (d) Obtain statistically valid results for (h), (i), and (j) above based on (2) (you can add the results from this regression to your table). Did this affect the substance of any of your conclusions (in other words, did the outcome of any of your tests change)?