# Homework

Introduction to Econometrics

Fall 2006

## Results

Using the 526 observations on workers available from Wooldrige's site in the Wage1 data set the following results are obtained. Consider the following linear model of log wages:

$$\ln(wage_i) = \beta_1 + \beta_2 educ_i + \beta_3 exper_i + \beta_4 exper^2 + \beta_5 female_i + u_i \tag{1}$$

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

### Least squares estimates

$$\widehat{\text{lwage}} = \underset{(3.820)}{0.390483} + \underset{(12.094)}{0.0841361}\,\text{educ} + \underset{(8.067)}{0.0389100}\,\text{exper} - \underset{(-6.389)}{0.000686022}\,\text{expersq}$$
$$- \underset{(-9.283)}{0.337187}\,\text{female}$$
$$T = 526 \quad \bar{R}^2 = 0.3950 \quad F(4, 521) = 86.685 \quad R^2 = 0.3996$$

($t$-statistics in parentheses)

1. $R^2$ is 0.3996.
2. Inclusion of *exper* and *expersq* allows the modeling of diminishing returns to experience. In this case, we expect $\beta_3 > 0$ and $\beta_4 < 0$.
3. $\partial E[ln(wage)]/\partial educ = \beta_2 = 0.0841$. Thus, an additional year of schooling results in an 8.41% increase in wages.
4. The RESET3 test statistic is: $F_{2,519} = 7.658522$, with p-value $= P(F_{2,519} > 7.65852) = 0.000527 < \alpha = 0.05$. The equation is misspecified at the 5% level of significance. Note, in practice you should also check the reset(2) statistic which is the t-ratio on $\hat{y}^2$ in an augmented model that contains this term. In this regression, the t-ratio is 3.77, which is significant at the 5 percent level.
5. Adding $educ^2$ to the model and rerunning RESET3 yields: Test statistic: F = 1.688653, with p-value $= P(F_{2,518} > 1.68865) = 0.186$. The null hypothesis of no misspecification is not rejected at the 5% level. The t-ratio for RESET2 is 1.44 and not significant. Things have certainly improved.
6. $\partial E[ln(wage)]/\partial educ = -0.0306529 + 2 * 0.00487164 * educ$. Therefore if Jethro has 6 years of schooling, the marginal effect is .0278. The marginal effect here is positive and so Jethro will benefit from finishing another year of schooling.
7. The estimated return to schooling for someone with a high school degree is marginal effect having 12 years of schooling is $-0.0306 + 2 * 0.00487 * 12 = 0.0862$. For a person with a college degree it is $-0.0306 + 2 * 0.00487 * 16 = 0.125$.

8. Test the null hypothesis that $\beta_5 = 0$ against the alternative $\beta_5 < 0$ using the t-ratio.

$$\widehat{\text{lwage}} = \underset{(5.323)}{1.01702} - \underset{(-1.005)}{0.0306529}\,\text{educ} + \underset{(3.862)}{0.00487164}\,\text{educ2} + \underset{(8.367)}{0.0398883}\,\text{exper}$$
$$- \underset{(-6.764)}{0.000719177}\,\text{expersq} - \underset{(-8.844)}{0.319578}\,\text{female}$$
$$T = 526 \quad \bar{R}^2 = 0.4107 \quad F(5, 520) = 74.182 \quad \hat{\sigma} = 0.40803$$
$$(t\text{-statistics in parentheses})$$

The t-ratio is $-8.84 < -1.645$ and the hypothesis of no discrimination is rejected at the 5% level.

9. The augmented model is

$$\begin{aligned} \ln{(wage_i)} =& \beta_1 + \beta_2 educ_i + \beta_3 educ^2 + \beta_4 exper_i \\ &+ \beta_5 exper^2 + \beta_6 female_i + u_i \end{aligned} \tag{2}$$

where $i = 1, 2, \ldots, 526$. The auxiliary model of the marginal effect of being female is

$$\beta_{6i} = \alpha_1 + \alpha_2 married_i \tag{3}$$

Substitution into the model (2) introduces and interaction term $female_i * married_i$ into the model. It's coefficient is

$$\begin{aligned} \ln{(wage_i)} =& \beta_1 + \beta_2 educ_i + \beta_3 educ^2 + \beta_4 exper_i + \beta_5 exper^2 \\ &+ \alpha_1 female_i + \alpha_2 female_i * married_i + u_i \end{aligned} \tag{4}$$

where $i = 1, 2, \ldots, 526$. The coefficient on $\alpha_2 = -2.215$ which is significantly negative at the 5% level (it is less than $-1.645$).

10. Reformulating the model to allow discrimination for married women only (not single women) amounts to omitting the $female_i$ in equation (4). This is probably **not** a good idea since $\alpha_1$ was significant. Also, the coefficients change quite a bit when $female$ is omitted–a likely sign off omitted variable bias.

11. The results table follows.

OLS estimates using the 526 observations. Dependent variable: lwage

| Variable | Model (a) | Model (e) | Model (i) | Model (j) |
|---|---|---|---|---|
| const | 0.390 | 1.017 | 0.968 | 0.8618 |
| | (0.102) | (0.191) | (0.191) | (0.195) |
| educ | 0.0841 | -0.0306 | -0.0278 | -0.0367 |
| | (0.00695) | (0.0305) | (0.0304) | (0.0312) |
| educ2 | - | 0.00487 | 0.00479 | 0.0053 |
| | - | (0.00126) | (0.00125) | (0.00128) |
| exper | 0.0389 | 0.0399 | 0.0426 | 0.0480 |
| | (0.00482) | (0.00476) | (0.00491) | (0.0049) |
| exper2 | -0.000686 | -0.000719 | -0.00076 | -0.000859 |
| | (0.000107) | (0.000106) | (0.000108) | (0.000109) |
| female | -0.337 | -0.319 | -0.255 | - |
| | (0.0363) | (0.0361) | (0.0461) | - |
| marrfem | - | - | 0.119 | -0.305 |
| | - | - | (0.0537) | (0.0430) |
| RESET2 (t-ratio) | 3.77* | 1.44 | | |
| RESET3 ($F_{2,\infty}$) | 7.6*5 | 1.689 | 1.447 | 0.912 |

(asymptotic standard errors in in parentheses)

Model (i) is the preferred model. All coefficients except the one on educ are significantly different from zero. There is evidence from RESET that $educ^2$ should be included in the model. Also the married/female interaction term is significant and should not be omitted. The $F_{2,\infty}$ 5% critical value is 3.00.