

EC 533 Labour Economics

Problem Set 1

Answers

1. Optimal Choice of Schooling and Returns to Schooling

(a) The individual maximizes

$$\begin{aligned} V &= \int_S^\infty ye^{-rt} dt = wf(S, A) \int_S^\infty e^{-rt} dt \\ &= \frac{w}{r} f(S, A) e^{-rS} \end{aligned}$$

or alternatively

$$\max \tilde{V} = \log w - \log r + \log f(S, A) - rS$$

The first order condition is

$$\frac{\partial \tilde{V}}{\partial S} = \frac{f_S}{f} - r = 0$$

Notice that

$$\log y = \log w + \log f(S, A)$$

so that

$$\frac{\partial \log y}{\partial S} = \frac{f_S}{f} = r$$

from the first order condition. The first order condition does not depend on w . This stems from the fact that w raises the cost of schooling and the additional earnings from schooling proportionately.

(b) First, derive the second order condition for the problem above

$$\frac{\partial^2 \tilde{V}}{\partial S^2} = \frac{f_{SS}f - (f_S)^2}{f^2} < 0$$

which implies, using the first order condition

$$f_{SS} - rf_S < 0.$$

Totally differentiate the first order condition

$$(f_{SA} - rf_A) dA = -(f_{SS} - rf_S) dS$$

so that we have

$$\frac{dS}{dA} = -\frac{f_{SA} - rf_A}{f_{SS} - rf_S}$$

This is positive iff

$$f_{SA} - rf_A > 0$$

For $f(S, A) = \phi(S)\delta(A)$ we have

$$f_{SA} = \phi'\delta' = \frac{\phi'\delta\phi\delta'}{\phi\delta} = \frac{f_S f_A}{f} = rf_A.$$

Notice that the basic human capital production function used by Griliches, $f(S, A) = \exp(\beta S + \gamma A)$ is multiplicative in S and A .

(c) S years of schooling now have an additional cost

$$C = \int_0^S ce^{-rt} dt = \frac{c}{r} (1 - e^{-rS})$$

and the maximization problem becomes after multiplying through by r

$$\max rV = wf(S, A)e^{-rS} - c(1 - e^{-rS})$$

1. The first order condition is

$$\frac{\partial rV}{\partial S} = (wf_S - rwf - cr)e^{-rS} = 0$$

or more intuitively

$$\underbrace{wf_S}_{\text{marginal benefit}} = r \underbrace{(wf + c)}_{\text{marginal cost}}$$

The second order condition has the form

$$w(f_{SS} - 2rf_S + r^2f) + cr^2 < 0$$

Multiplying the FOC by r and substituting into the SOC yields the simplification

$$f_{SS} - rf_S < 0$$

which is the same as in the original problem. Totally differentiating the first order condition

$$w(f_{SS} - rf_S)dS + (f_S - rf)dw - (wf + c)dr - rdc = 0$$

results in the following comparative statics results

$$\frac{dS}{dr} = \frac{wf + c}{w(f_{SS} - rf_S)} < 0$$

$$\frac{dS}{dc} = \frac{r}{w(f_{SS} - rf_S)} < 0$$

$$\frac{dS}{dw} = -\frac{f_S - rf}{w(f_{SS} - rf_S)} > 0$$

The last inequality follows from the FOC

$$f_S - rf = \frac{rc}{w} > 0$$

2. An increase in the wage now raises optimal schooling. An increase in the wage still raises earnings proportionately but affects costs less than proportionately because foregone wages are only part of the total costs. In order to balance marginal costs and benefits again after an increase in w schooling will have to be raised to lower the marginal benefit and raise the marginal cost.
 3. By totally differentiating the first order condition $wf_S - wrf = cr$ it is easy to see that it is still true that $dS/dA > 0$ if $f_{SA} - rf_A > 0$. Using $f(S, A) = \exp(\beta S + \gamma A)$, this condition becomes $f_{SA} - rf_A = (\beta - r)\gamma f = \gamma rc/w$, where the last equality follows from the first order condition. It is easy to see that this expression has the sign of c if $\gamma > 0$.
- (d) The human capital production function implies

$$\frac{f_S}{f} = AbS^{b-1} = r$$

Notice that this can be rewritten as

$$\frac{rS}{b} = AS^b = \log f(S, A)$$

so that

$$\log y = \log w + \frac{r}{b}S$$

Since S is the only random variable on the right hand side it follows immediately that

$$\text{cov}(\log y, S) = \frac{r}{b}\text{var}(S)$$

so that OLS overestimates the returns to schooling because of ability bias.

2. Estimating Returns to Schooling

- (a) We know that the first order condition for this model is

$$\frac{\partial \log y}{\partial S_i} = \beta_1 + 2\beta_2 S_i = r$$

so that

$$S_i^* = \frac{r - \beta_1}{2\beta_2}.$$

We cannot estimate the returns to schooling in this case because the choice of schooling is the same for everybody, so there would be no variation in the regressors.

- (b) If the interest rate is random, schooling choices now become

$$S_i^* = \frac{r_i - \beta_1}{2\beta_2}.$$

The coefficients of a regression of log earnings on schooling and it's square are

$$\begin{aligned}\text{plim } \hat{\beta}_1 &= \frac{\text{var}(S_i^2)\text{cov}(\log y, S_i) - \text{cov}(S_i, S_i^2)\text{cov}(\log y, S_i^2)}{\text{var}(S_i)\text{var}(S_i^2) - \text{cov}(S_i, S_i^2)^2} \\ \text{plim } \hat{\beta}_2 &= \frac{\text{var}(S_i)\text{cov}(\log y, S_i^2) - \text{cov}(S_i, S_i^2)\text{cov}(\log y, S_i)}{\text{var}(S_i)\text{var}(S_i^2) - \text{cov}(S_i, S_i^2)^2}.\end{aligned}$$

The key covariances in these expressions are

$$\begin{aligned}\text{cov}(\log y, S_i) &= \beta_1\text{var}(S_i) + \beta_2\text{cov}(S_i, S_i^2) \\ \text{cov}(\log y, S_i^2) &= \beta_1\text{cov}(S_i, S_i^2) + \beta_2\text{var}(S_i^2)\end{aligned}$$

so that

$$\begin{aligned}\text{plim } \hat{\beta}_1 &= \frac{\text{var}(S_i^2) [\beta_1\text{var}(S_i) + \beta_2\text{cov}(S_i, S_i^2)] - \text{cov}(S_i, S_i^2) [\beta_1\text{cov}(S_i, S_i^2) + \beta_2\text{var}(S_i^2)]}{\text{var}(S_i)\text{var}(S_i^2) - \text{cov}(S_i, S_i^2)^2} \\ &= \frac{\beta_1 [\text{var}(S_i)\text{var}(S_i^2) - \text{cov}(S_i, S_i^2)^2]}{\text{var}(S_i)\text{var}(S_i^2) - \text{cov}(S_i, S_i^2)^2} \\ &\quad + \frac{\beta_2 [\text{var}(S_i^2)\text{cov}(S_i, S_i^2) - \text{var}(S_i^2)\text{cov}(S_i, S_i^2)]}{\text{var}(S_i)\text{var}(S_i^2) - \text{cov}(S_i, S_i^2)^2} \\ &= \beta_1\end{aligned}$$

and analogously

$$\text{plim } \hat{\beta}_2 = \beta_2.$$

There is no ability bias here because the human capital production function is multiplicative in S_i and ε_i (see question 1 (b)).

- (c) To make returns heterogeneous, make β_1 a function of ε_i :

$$\beta_1 = \theta_0 + \theta_1\varepsilon_i.$$

Substituting in the first order condition yields

$$S_i^* = \frac{r - \theta_0 - \theta_1\varepsilon_i}{2\beta_2}.$$

The covariance is

$$\text{cov}(S_i^*, \varepsilon_i) = -\frac{\theta_1}{2\beta_2}\sigma_\varepsilon^2 > 0$$

if $\theta_1 > 0$. Whether more able individuals have higher or lower returns to schooling depends on whether we believe that schooling and ability are complements or substitutes. Some schooling may make up for lack of ability, but we probably believe that ability also helps going further in school (after all, that's why we are all going through graduate training, isn't it?).

- (d) Note that the optimal schooling level is from part (a)

$$S_i^* = \frac{(1 - z_i)r_0 + z_i r_1 + \eta_i - \beta_1}{2\beta_2}.$$

The probability limit of the Wald estimator is

$$\begin{aligned}
\text{plim } \hat{\beta}_W &= \frac{E(\log y_i | z_i = 1) - E(\log y_i | z_i = 0)}{E(S_i | z_i = 1) - E(S_i | z_i = 0)} \\
&= \frac{E(\beta_0 + \beta_1 S_i + \beta_2 S_i^2 + \varepsilon_i | z_i = 1) - E(\beta_0 + \beta_1 S_i + \beta_2 S_i^2 + \varepsilon_i | z_i = 0)}{E(S_i | z_i = 1) - E(S_i | z_i = 0)} \\
&= \frac{\beta_1 [E(S_i | z_i = 1) - E(S_i | z_i = 0)] + \beta_2 [E(S_i^2 | z_i = 1) - E(S_i^2 | z_i = 0)]}{E(S_i | z_i = 1) - E(S_i | z_i = 0)} \\
&= \beta_1 + \beta_2 \frac{E(S_i^2 | z_i = 1) - E(S_i^2 | z_i = 0)}{E(S_i | z_i = 1) - E(S_i | z_i = 0)} \\
&= \beta_1 + \beta_2 \frac{\frac{r_1^2 - 2r_1\beta_1 - r_0^2 + 2r_0\beta_1}{4\beta_2^2}}{\frac{r_1 - r_0}{2\beta_2}} = \beta_1 + \frac{1}{2} \frac{(r_1 - r_0)(r_1 + r_0 - 2\beta_1)}{r_1 - r_0} \\
&= \frac{1}{2}(r_1 + r_0).
\end{aligned}$$

The population average return to schooling is

$$\begin{aligned}
E\left(\frac{d \log y_i}{d S_i}\right) &= E(\beta_1 + 2\beta_2 S_i) \\
&= \beta_1 + 2\beta_2 E\left(\frac{(1 - z_i)r_0 + z_i r_1 + \eta_i - \beta_1}{2\beta_2}\right) \\
&= \beta_1 + 2\beta_2 \frac{(1 - P(z_i = 1))r_0 + P(z_i = 1)r_1 - \beta_1}{2\beta_2} \\
&= (1 - P(z_i = 1))r_0 + P(z_i = 1)r_1.
\end{aligned}$$

The return to schooling for an individual is equal to the interest rate faced by each individual, and on average these returns are r_0 and r_1 for each of the respective groups. The population return is a weighted average of the two returns in the population. The population weights depend on the fraction of the population faced with relatively low and high interest rates. The Wald estimator also estimates an average of the population returns, but the weights here are fixed and don't depend on the population composition (because we are simply comparing the means for the low and high interest rate groups, no matter what the size of the groups is). Thus, only if $P(z_i = 1) = \frac{1}{2}$ do the weights coincide and we get

$$E\left(\frac{d \log y_i}{d S_i}\right) = \frac{1}{2}(r_1 + r_0)$$

which is equal to the probability limit of the Wald estimator.