Steve Pischke

EC 533 Labour Economics
Problem Set 1
Answers

1. Optimal Choice of Schooling and Returns to Schooling

(a) The individual maximizes

$$V = \int_S^\infty y e^{-rt} dt = w f(S,A) \int_S^\infty e^{-rt} dt$$
$$= \frac{w}{r} f(S,A) e^{-rS}$$

or alternatively

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

The first order condition is

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

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

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

which implies, using the first order condition

$$f_{SS} - r f_S < 0.$$

Totally differentiate the first order condition

$$(f_{SA} - r f_A)\, dA = -\left(f_{SS} - r f_S\right) dS$$

so that we have

$$\frac{dS}{dA} = -\frac{f_{SA} - r f_A}{f_{SS} - r f_S}$$

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

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

and the maximization problem becomes after multiplying through by r

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

1. The first order condition is

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

or more intuitively

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

The second order condition has the form

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

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. Using $f(S,A) = \exp(\beta S + \gamma A)$, note that $f_s = \beta f$, so that the first order condition becomes

$$\begin{aligned} wf_S - wrf &= cr \\ w(\beta - r) f &= cr \end{aligned}$$

and the second order condition requires $r > \beta$. Hence

$$\begin{aligned} w\frac{r-\beta}{r}\exp(\beta S + \gamma A) &= -c \\ \ln(w) + \ln\left(\frac{r-\beta}{r}\right) + \beta S + \gamma A &= \ln(-c) \\ S &= \frac{1}{\beta}\left[\ln(-c) - \ln(w) - \ln\left(\frac{r-\beta}{r}\right) - \gamma A\right] \end{aligned}$$

and $dS/dA = -\gamma/\beta < 0$.

(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

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

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

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

The key covariances in these expressions are

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

so that

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

and analogously

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

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

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

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

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

The population average return to schooling is

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

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

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

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

3. Subject choice

(a) To solve this model, find the optimal S_s for each subject choice, and then pick the choice which gives the higher PDV. The value of schooling for subject choice s is identical to the standard model:

$$
\begin{aligned}
V_s &= \int_{S_s}^{\infty} w_s f(S_s, A) e^{-rt} dt \\
&= \frac{w_s}{r} f(S_s, A) \left[e^{-rS_s} \right]
\end{aligned}
$$

Because the FOC

$$\frac{f_{S_s}}{f} = r$$

is independent of w_s we get immediately

$$S_1 = S_2$$

As a result, V_s only differs by subject because wages w_s differs by subject. As a result, inviduals will choose the subject with the higher w_s.

(b) In the model, all individuals choose one subject, and nobody studies the other one. This is not very interesting as a model of subject choice as we would like to have some variation in subject choice.
There are various ways of fixing this problem. One would be to make the $f(.,.)$ function subject specific, so that ability matters differentially for the subject specific returns. Another one would be to introduce a cost of schooling which differs across indivdiuals and by subject. In either of these cases there will be a paramter (the relative return to ability by subject or the relative cost parameter for each subject) which differs across individuals. The equilibrium choices will feature a cutoff for these distributions so that individuals with a particularly high ability/low cost for a subject choose this particular subject. In the case of a schooling cost, this will also imply that the choice of the number of years of schooling S_s is now subject specific ($S_1 \neq S_2$ for an individual). As a result, there will be a distribution of choices of years of schooling which will depend on the underlying parameter distributions in the population.

(c) There are two challenges. One is that subject choice may be subject to ability bias just as there is ability bias in the length of schooling. As long as subject choice is only driven by costs this may not be a problem but it becomes relevant if ability is related to subject choice.
The second challenge is the fact that subject choice and the length of schooling may interact. Suppose individuals in the comparatively low return subject tend to choose fewer years of schooling. If we were to study subject choices in college, then some of these individuals may have chosen not to go to college at all. As a result, the individuals we do actually observe in low return subjects in college may be particularly high ability individuals and we would tend to overestimate the return to subject choice among college students.

(d) The IV assumptions are first stage, quasi-random assignment, and the exclusion restriction. To the degree that peers matter and girls like science less, more girls may influence everybody to be less interested in science. So there is potential for a first stage. Random assigment depends on the setting–is there selection in who is in a particular high school class (e.g. are the schools with more girls maybe higher ability?) or does variation come about because students are randomly allocated to classes and the fraction girls will vary for random reasons in small classes? Exclusion is the most dubious assumption here. Girls could influence not just subject choice but may influence other academic outcomes. Or more girls might mean that boys are more likely to find a girl friend who distracts them from academic matters. Finally, girls may affect selection into who goes to college and this may differ by subject choice; the same selection problem as discussed in part (c).