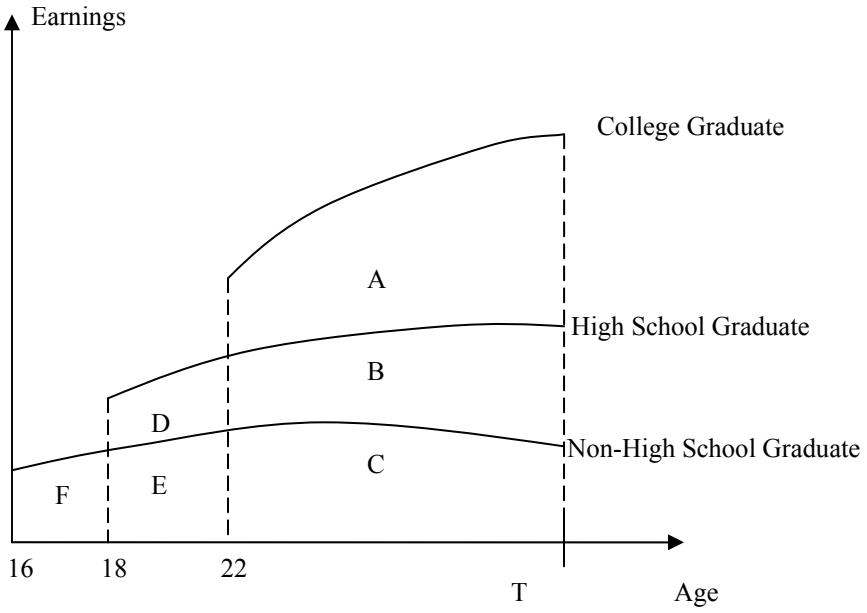
Human Capital Theory

A short introduction to Economics of Education: Human Capital Theory

Human Capital Theory says that

- → Investments are made in human resources so as to improve individual productivity and therefore their earnings.
- → It is an investment because costs are incurred, both explicit, in the form of fees, and implicit, in the form of "hardship", as well as opportunity cost of staying in school while the individual could begin his career.
- → Like most of studies in Economics, the optimal choice is dependent on the balance between benefits and costs.
- → There is an additional element in the study of this cost and benefit of investment in human capital, that of private versus social cost and benefit. As the terms imply, the former pertain to the individual agent's personal cost in terms of hardship (which is a function of their innate talent), and opportunity cost (which is a function of personal network, and opportunities available at any given time). The latter pertains to the optimal choice if there was one individual who could make a binding choice for everyone of us.

What does this theory predict?



Lifetime Income Stream by Educational Attainment

Basic Model

Assumption:

1. Individual does not receive any direct utility or disutility from the educational process.

2. Hours of work are fixed, including time spent in school.
3. Income streams in all states are known with certainty.
4. Individuals can borrow and lend at the real interest rate r.

This simplifies the model to a comparison between total benefit in terms of income stream, and cost of education. That is we do not model the hardship element in staying in school. Further note that agents here are homogenous, i.e. one agent is no different from another. Or at least, we are comparing one type of agent at a time. Why? Well, since all we are concerned with is the mechanism that directs an individual's choice, there is no gain in adding that additional level of complication. To consider individuals of other types, who might find higher educational attainment easier, or another who might find it more difficult, may easily be factored by including a constant term to the direct cost element.

How do we proceed then?
Consider first an individual who chooses to drop out of high school at age 16, and starts his career. His lifetime income stream is just the area of F+E+C, or

$$\sum_{t=16}^{T} \beta^{t-16} Y_t^{drop}$$

where β is the discount factor we have talked about earlier.

However, if he completes his high school education and he lifetime income stream is B+C+D+E, or

$$\sum_{t=18}^{T} \beta^{t-18} Y_t^{high}$$

However, he faces the cost of not quitting, and staying in school. Let the cost of school fees be F_t^E where E is "*high*" for high school graduate, and "*college*" is for college graduate. Then his net benefit from staying in school is,

$$\sum_{t=18}^{T} \beta^{t-18} Y_t^{high} - \sum_{t=16}^{17} D_t^{high} - \sum_{t=16}^{T} \beta^{t-16} Y_t^{drop}$$

He would choose to stay in school and complete his high school diploma if and only if the net benefit is positive.

At graduation from high school, he faces a similar problem, one that might be similar to your considerations when you choose to start your undergraduate stint in StFX. His projected income stream and net benefit are

$$\sum_{t=22}^{T} \beta^{t-22} Y_t^{college}$$

and

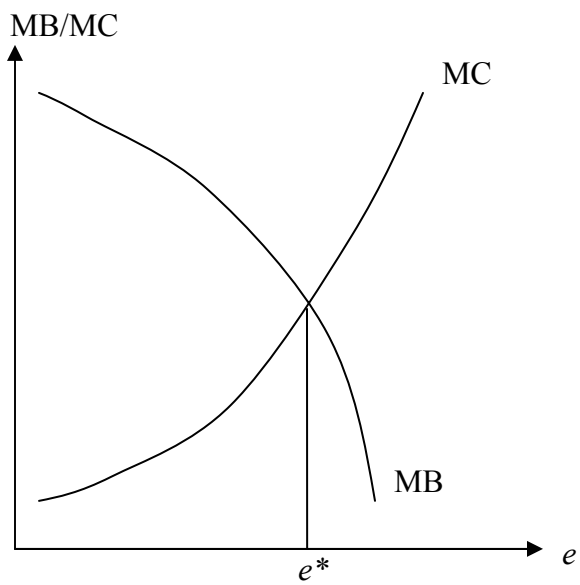
$$\sum_{t=22}^{T} \beta^{t-22} Y_t^{college} - \sum_{18}^{21} D_t^{college} - \sum_{t=18}^{T} \beta^{t-18} Y_t^{high}$$

respectively. The individual would choose to pursue his college degree if and only if the net benefit is greater than zero. Note that we have assumed that $Y_t^{college} > Y_t^{high} > Y_t^{drop}$

Of course if we extend our argument beyond a single type of individual, it just means that the net benefit for a high type individual is greater by virtue of perhaps a steeper, and higher income profile after college. Of course, that is on the assumption he does not embark on the PhD route! Why? If having a PhD increases the income profile substantially, why don't we see more individuals choosing to pursue PhD?

Another Model

Can we model this mechanism in any other manner? Consider the following general model. Let the benefit of staying in school be $B(e, \beta \mid p)$, where e is the number of years of education, and p is the type of individual. Where B is increasing and concave in e, which implies that although more years of education is better for an individual, the marginal gains from an additional year eventually starts falling. The cost of staying in school is $C(e, \beta \mid p)$ where C is increasing and convex in e. This just says that the cost of an additional year of school will rise very faster with each additional year, reflecting the greater opportunity cost, since there is an implicit threshold the labor market is willing to pay a rookie. Then equilibrium choice occurs when $B_e = C_e$ and diagrammatically,



Is it as simple as the above?

1. This simple model cannot tell us why some people pursue education beyond the standard Bachelors degree, or Professional Degrees. How can we model that question? The model could be augmented by an enjoyment variable, so that individual's more inclined to academic learning or inventing or research gains utility with every additional year of school. While others less inclined in those

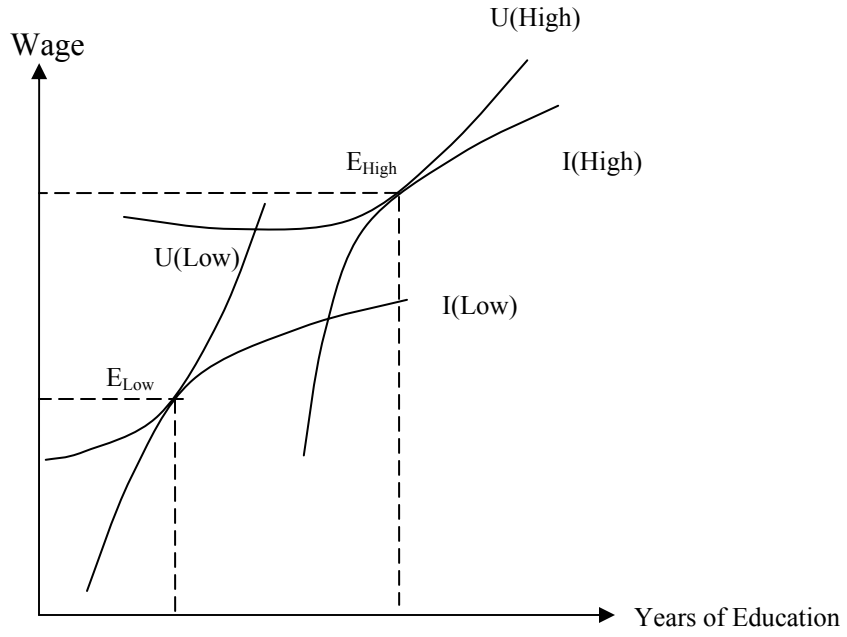
endeavors suffer disutility from staying an additional year in school. Note that such a model may also explain individuals who select into differing grades or levels of expertise. For example, you have carpenters who enroll in trade schools, or simply learn on the job, and you have individuals more inclined towards carpentry as an art form, and undertake study or apprenticeships with Masters craftmen. This consideration hence attempts to model differential in schooling choices by arguing that individuals differ in their marginal rate of return to schooling. This differential could just as easily be modeled by assuming different individuals have different discount rates β.

2. The model also ignores the difficulty in borrowing since college education does not come cheap these days, and your future earnings is not a sufficiently good collateral for the financial institutions. This is further complicated when we consider the fact that an individuals income stream is in fact stochastic, i.e. not known with certainty.

Education and Market Equilibrium

In a labor market, differing firms required labor of differing quality and skill, and hence differing levels of education. Note that in this depiction, as the isoprofit moves towards the south east, they are on higher isoprofit curves, since they are paying lower wages, and yet getting labor of increasing skills. An example of this differentiation is between the requirements for a Corporate Banker, and a subway train driver.

- → Note the difference in the marginal rate of substitution between a high type individual who is more inclined to have more education, and low type individuals who are not. For the latter type, a larger wage increase is required to entice them into greater number of years of education.
- → Note also that high type firms are more inclined / prepared to paying higher wages for the additional years of education.



→ In a competitive equilibrium, firms earn zero profit. This then mean that the "hull" formed by the isoprofit curves constitutes the employers' offer curve, showing the maximum wages that can be paid for various levels of education. Note the following points;

- Labor with more education enjoy a compensating differential (for the difference in the amount of years of education). This is dependent on the
  - preferences of different types of workers,
  - the distribution of these different types of workers, which determines the firms' choice of staff, and
  - the technology used in production.
- Employers with the different preferences for differing types of employees, are matched with those employees.

→ This equilibrium is pareto optimal. Note an important caveat: This assumes education is fully revealing, and representative of the abilities a firm needs, i.e. there is **no asymmetry in information**.

So what do we want to verify with data:

1. **Estimating the returns to Schooling**

   The typical method of estimating the rate of return to schooling uses data on the earnings (such as annual income) and schooling of different workers and estimates the percentage wage difference associated with one more year of schooling, after "controlling" or accounting for worker characteristics such as race, sex and age. The typical regression equation is

$$\log w_i = \alpha s_i + \mathrm{A}X_i + \varepsilon_i$$

Where w is the income, and X is all other characteristics that describes the individual. Then α tells you the percentage change in wages for a change of a year in education. Why this interpretation? Can you show this?

What are some problems with this regression:
The comparison is valid if and only if all individuals are the same, except for the choices they make. But our priors suggests that this may not be true. If we all differ on some level, let it be "ability" that is perhaps genetically transmitted, then the error in our estimation suffers from endogeniety, i.e. our "ability" in the errors is systematically correlated with our schooling variable s. So what's the big deal then? Let a be our ability, such that the relationship between a and s is as follows:

$$s_i = a_i + \upsilon_i$$

and what we are really interested in is

$$\log w_i = \alpha s_i + \lambda a_i + \mathrm{A}X_i + \varepsilon_i$$

But because a is not observable, our actual regression is actually:

$$\log w_i = \alpha s_i + \lambda(s_i - \upsilon_i) + \mathrm{A}X_i + \varepsilon_i$$
$$\Rightarrow \log w_i = (\alpha + \lambda)s_i + \mathrm{A}X_i - \lambda\upsilon_i + \varepsilon_i$$

So what's the big deal?

a. First note that the coefficient of s measure now not just the impact of schooling, but ability as well. Which means we cannot discern the true impact of schooling.

b. The form of the errors violate the assumption that errors are normally distributed with mean 0.

Remedy:

1. Use a proxy for ability such as IQ measure. However, this is highly contentious since what does IQ really measure even in the context of academic achievement. And how does it translate to work ability.

2. We could always compare apples to apples. That is use a "Natural Experiment", where we perform the above regression on twins. Each twin would then have similar ability. How does it work? Take a pair of twins, run the similar regression by on the difference in income against difference in schooling. Because they were born of same parents, ability "*a*" will get cancelled out. And all our assumptions are met. Or does it? The results has been mixed (See Borjas). But why when the method seems so appealing and intuitive? Well consider this, if their abilities are the same, why do they differ in their choice but because they differ in their discount rates. But why do they differ in discount rates, but because there is still something unobservable, and we are back to square one.

3. Angrist and Krueger (1992) (Read Borjas page 251 for reference) estimated the above equation using Instrumental Variables regression, using the natural experiment that the Vietnam war created. The estimate they got was that an additional year of schooling raised income by 7%. Another example of using Instrumental Variables was also performed by Angrist and Krueger (1991) (Again read Borjas for specifics) where they used compulsory schooling which

effectively forced everyone to stay in school regardless of ability until a specific age, which depends on the year we are considering, and state the sample was from. Then individuals who remain past the law are of one ability, and those who dropout are of another.

2. **If schooling is that important, and may level the playing field, why don't we spend more money on the public schooling system since their may be social gains.**

   The next question then is whether dumping money into school system enhance the quality of our schools, and hence raising the rate of return to schooling? Card and Krueger (1992) found evidence that school quality was positively correlated with rate of return to schooling. Borjas gives a summary in section 7.7 in his book. There has also been mixed results in this area, see particularly Hanushek.

3. **Do workers really maximize lifetime earnings?**

   Doesn't point 1 answer this question? Well, not exactly. Point 1 assumes that our basic model is correct, we do not know the true mechanism that goes on within the mind of the individual unless we can see how he reacts under differing circumstances, in this case offerings of differing lifetime income stream. But we cannot perform this ideal experiment, since no individual can rescind a choice he has made. Even if he did, he would have been at a different age from when his initial decision was observed. What this means is that even if we could track an individual's choice, the choice the individual makes is dependent as before on both observables and unobservables, and their will be selection bias since an individual of any particular "type" who make their choice conditional on his "type". The technique used is generally called "Selection Bias Correction" (2 stage estimation methods. To see more details refer to Borjas.). Results using these techniques has found that our theory is correct, i.e. that individuals make their choices conditioning on their ability or "type", and there is hence no one "type" of individual. Then the differential in wages paid is a culmination of differential in demand, and length of schooling required for those task for which individuals are hired. Or does it?

Education as a Filter

Our previous study has assumed that education raises a worker's productivity, and hence the warrants higher wages the longer an individual stays in school. We now an alternative where education does not augment an individuals skill set, but is just a sorting mechanism for individuals of greater levels of "intelligence" from individuals with a lower endowment of it. This works if and only if a potential employer cannot discern between the true type of the individual, i.e. that is asymmetry of information between employers (referred to as Principals) and workers (agents).

A sketch of the model:

1. Let there be two types of workers, indexed by *h* for high type, and *l* for low type. Let there be a unit mass of all workers, and that of this *q* are of the former type, and 1-*q* are of the latter. The former commands a higher lifetime stream of income, y_h, and the latter commands y_l.
2. If all employers could discern between all the worker types, they would offer each type their respective incomes. Thus the justification for our assumption of asymmetry on information between the agents and the principals. The idea is as follows: How can we sort this different types of workers such that they would reveal themselves, since there is no incentive for the *l* type individuals to reveal themselves to you. Put another way, *l* type individuals would be inclined to lie so that he gains the higher wage rate or income. Assuming no punishment for lying, and that it takes time for employers to discover an individuals true type.
3. One possible choice if the principal cannot discern between the types, and barring any other information, the principal would choose to offer a weighted or average income for both *h* and *l* type. Then the income is $y = qy_h + (1-q)y_l$. So what are some implications?

a. *l* type individuals would prefer this outcome since they gain a higher level of income.
b. Neither the principal nor the *h* type individual would like this since *h* type are now paid less, and the principal would find that he is mismatching worker task to type, and the firm would be working at sub-optimal efficiency.

4. What can be done? Well, if there is a form of signal that would provide the separation that is required, we would be able to offer the right wage to the correct type. Here's where number of years of schooling comes in. Let the cost of the number of years of schooling be a constant g_t where *t* indexes the individuals type. Further, we need to assume that the *h* type individuals have a lower cost to schooling. What the principal wants is to find the number of years of schooling which the *l* type would find not beneficial to him given his type, and he would consequently reveal his type by the amount of schooling he chooses.

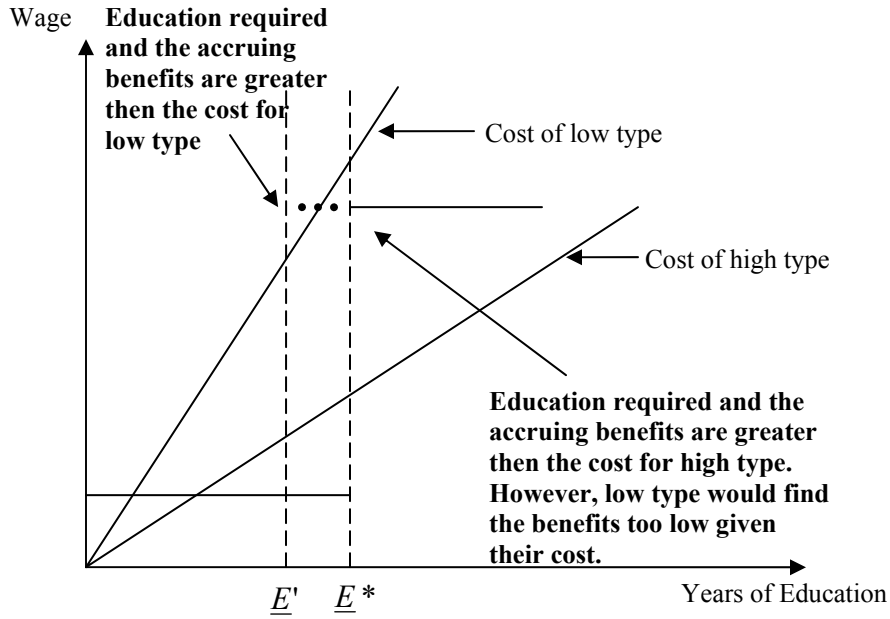
$$y_l > y_h - (g_l \times y^R)$$
$$\Rightarrow y^R > \frac{y_h - y_l}{g_l}$$

Where y^R is the minimum or reservation level of education that the principal must use to induce revelation of the individual's type. Note that the greater the differential between the incomes offered to the two types, the greater would the minimum schooling has to be. However this is only part of the story. The principal would have to ensure that the *h* type individual would actually choose to exceed the requirements. That is

$$y_l < y_h - (g_h \times y^R)$$
$$\Rightarrow y^R < \frac{y_h - y_l}{g_h}$$

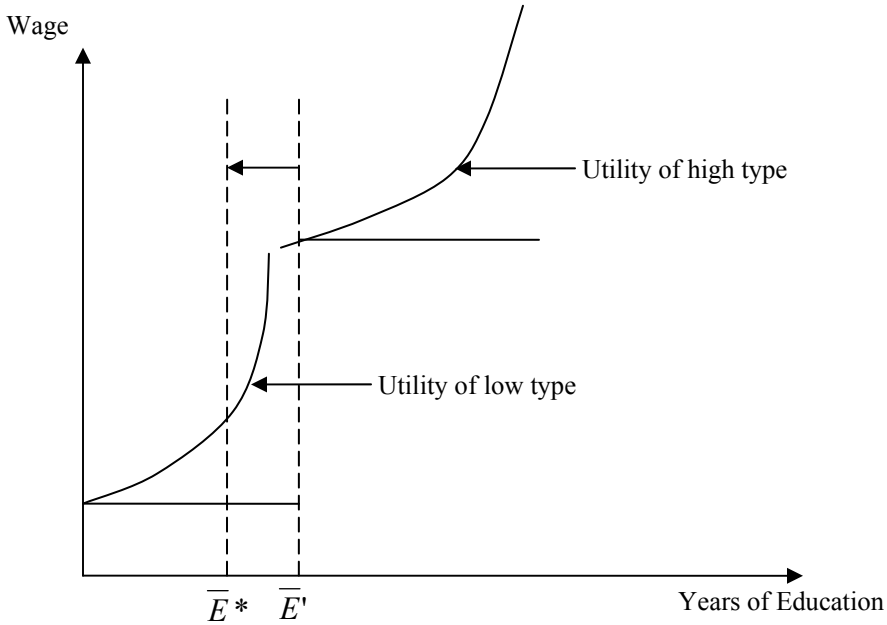
And this is the solution.

Does reality conform to the basic model we had or this one?



Diagrammatically, this simple model is illustrated above, where $\underline{E}*$ is the efficient education cutoff choice. Note that at $\underline{E}*$ the low type are better off revealing their true type then lying about their capabilities. While at $\underline{E}'$ the benefits in the form of wages are high enough such that low type would find it beneficial to lie.

The above model can also be considered from the point of view where different individuals have different preferences towards education, and hence the amount of compensation they require to be on a higher indifference curve.



On Job Training and the Age Earnings Profile

We have thus far considered human capital investment from the perspective of education. However, job training is not pertinent to youths prior to entry into the labor force. As technology evolves labor has to constantly retrain to maintain there relevancy.

There are two types of On-Job-Training:

1. **General Training**: This form of training is not task specific, and enhances the value of the worker anywhere with any firm he finds himself with. Example: Programming skill in Java, HTML, and C+ and C++ are transferable skills in any firm requiring a programmer.
2. **Specific Training**: This form of training are firm and/or task specific, and may be relevant to a firm or industry only, i.e. it is not an easily transferable skill. Example: Learning how to use SAP is useless if the firm uses are client software in monitoring their production, and inventory.

Consider a simple model where a firm is contractually hired for only two periods, after which the worker retires. The firm of course continues to exist. Let it be such that the firm only hires new labor after retiring the current worker. Hiring is instantaneous, and there is no lag. The profit maximizing choice of this firm is then derived from the following problem;

$$\max_L \pi = [P_1 F(L) - C(L)] + \frac{[P_2 F(L) - C(L)]}{1+r}$$

Then the optimal level of employment is given by

$$P_1 MP_L + \frac{P_2 MP_L}{1+r} = VMP_1 + \frac{VMP_2}{1+r} = MC_L + \frac{MC_L}{1+r} = C_L + \frac{C_L}{1+r}$$

If however the firm needs to train all employees in the first period while they are on the job, the condition needs to be augmented with this cost, as well as the fact that the employees are more productive in the second period after training. Let the marginal cost of training be O and that this cost is incurred only in period 1, and the marginal cost of labor is just wages. Further, let the labor market be competitive, so that after training, the worker with a higher skill may sort employment elsewhere if the wage offer is sufficiently high, that is the wages may be different. The condition is now;

$$P_1 MP_{L,NT} + \frac{P_2 MP_{L,T}}{1+r} = w_1 + O + \frac{w_2}{1+r}$$

Who pays for training:

1. General Training: As noted, general training is transferable. This then mean that the skills are beneficial to both the employee, and the employer. Let's use backward induction:
    a. In the second period, because the skills are perfectly transferable, let the market wage be w_2, then the firm has to pay the worker this competitive wage which is equal to the value of the marginal product of labor in the second period, else because the skill is transferable, they would choose to leave if a better offer arrives.
    b. However, in the first period by the optimality condition, it reduces to

$$P_1 MP_{L,NT} = w_1 + O$$
$$\Rightarrow P_1 MP_{L,NT} - O = w_1$$

That is the first period wage offered is paid for by the employees themselves, i.e. the cost of training is completely transferred to the employees.

2. What if the skills are specific? The cost in this instance will be shared. The reason is as follows. If the firm pays for the training, it would have to recover the cost in the second period by paying below the value of marginal product of labor post training, so that the worker earns w_1 or the going wage in the market. This works if the cost of training is not high, so that what the worker might get in the market is still below what is being paid to him in the second period. HoweverIf the training cost is high such that $VMP_1 < O + w_1$ However, if the cost of training is sufficiently low because the training is specific
   a. Case 1: Consider the firm paying the cost of training in the first period, *O*, and reaping the rewards by paying the workers a wage below their second period value of marginal product, $w_2 < VMP_2$. Then the firm stands to gain if and only if the worker stays with the firm. However, if the worker were to leave, they would be stuck with a cost of O which they had accrued in the first period. Consequently this strategy is not viable. (Is the threat of leaving by the worker credible as long as the wage offered is slightly above the going wage should he leave the firm?)
   b. Case 2: But what if the firm makes the worker pay for the training. The worker would then incur a lower actual wage in the first period, and reap the rewards in the second period. However, the firm can threaten to renege on the deal, and pay below the value of marginal product in the second period, or even dismiss the worker. (The fear of being dismissed is in truth not a problem since the firm cannot credibly commit to making that move since it can always make a positive gain by retaining the worker because the value of marginal product is greater.)
   c. Case 3: The hints in brackets give in essence the answer. If the two share in the cost and benefits, the threat of quitting and dismissal becomes not credible. Suppose the general skill of the worker in each period warrants a wage of *w*. Then the firm can simply pay a wage higher then that, say w^R, but lower than the value of marginal product of labor in the second period. In such a case, the worker would never consider quitting. So the firm's share of the gain is $\frac{VMP_2 - (O + w^R)}{VMP_2 - VMP_1}$ while the workers share is just $\frac{w^R - VMP_1}{VMP_2 - VMP_1}$. Note that this means in equilibrium, the workers are paid a wage lower than the going wage while training, and a higher than going wage after training. Your text has a diagrammatic representation on page 276.

Mincer Earnings Function

Our examination of human capital model culminates in a age-earning profile for individuals. Jacob Mincer in fact showed that the human capital model generates an age-earning profile of the form

$$\log w = as + bt - ct^2 + DX$$

where w is the worker's wage, s is the number of years of schooling, t is the number of years in the labor market (experience), and X is the set of other pertinent variables. The equation above is typically referred to as the Mincer Earnings Function. The regression coefficients has the following interpretation; a is the percentage change in wages for an additional year of schooling, and t is the percentage change in wages due to an additional year of experience. The latter is typically interpreted as the impact of on-job-training.