# Human Capital: First Lecture

LABOR ECONOMICS (ECON 385)

BEN VAN KAMMEN, PHD

# Introduction

Compensating wage differentials reveal that jobs are different—and that employees are different with respect to their preferences for job attributes. Consequently wages vary because of differences among jobs. Wages vary further because the people doing them have different levels and varieties of human capital—the set of skills and abilities that augment workers' levels of productivity.

- Human capital includes schooling and formal education programs. Also on-the-job training and apprenticeships are examples, as well as learning by doing that occurs as a person performs his job over time.
- Human capital can be general in nature (useable on many jobs) or specific to a single job or firm.
- Individuals choose whether to acquire training based on a calculation of present discounted benefits and costs.
- Models of human capital investment discussed in this class treat the subject a lot like the financial investment example above.

# General and specific human capital

- Examples of general human capital:  keyboarding, arithmetic, teamwork, and using email.
- Examples of job specific human capital:  layout and assembly skills used by a carpenter, welding, bartending, and purchasing.
- Examples of firm specific human capital:  knowledge of organizational structure, knowledge of software designed by/for one's employer, familiarity with the layout employer's facilities, where to find office supplies, etc., how to use a machine that the firm has patented.
- With some types of training (firm specific), the employer has more of an incentive to invest in provision of the training than others (general).

# Present value calculations

- "Present Discounted" refers to the fact that many of the benefits from human capital occur in the future.
- Agents are presumed to value present and future costs/benefits differently—namely that they value benefits they expect in the future less than benefits received in the present.
- Discounting future benefits appropriately enables the economic agent to compare costs and benefits that are equivalent to "as if" they were received today.
  - Example: an individual is indifferent to having $10 today and $10.50 tomorrow. This person's temporal discount rate is the solution to the following algebra problem:

$$\frac{10}{x} = 10.5 \rightarrow x = \frac{1}{1.05} \approx 0.952.$$

# Present value calculations (continued)

- So with this discount rate in mind, the agent would compare all costs and benefits in "present discounted" terms. E.g., if he has to decide whether to buy a riskless asset for $50 today that will earn him $52 tomorrow, he makes the following calculation:

$$\text{(Net)Present Discounted Value } (PV) = \frac{52}{1.05} - 50; \text{buy if } PV \text{ is } > 0.$$

$$PV = -0.476$$

So even though this asset is "profitable", it isn't worth it for this person to delay his consumption until tomorrow.

- The larger is the temporal discount rate, the more likely an agent will prefer benefits received today to benefits received in the future. Agents with large discount rates are characterized as "impatient".

# Costs and benefits of acquiring human capital

- To acquire human capital, there is a time cost (paid in the present) and usually a monetary cost (paid in the present or with a loan whose payments can easily be discounted into present value terms). Once the skills are acquired the worker's productivity is increased, along with his wage, and he earns a higher wage in all future periods until he "retires".
- So the question is: "is it worth it to sacrifice some time and money today to get more money later?"

# Human capital model

To solve an individual's human capital accumulation question, we need to consider how much he stands to gain by investing and what it will cost. Consider a model of life time earnings with the following assumptions:

- An agent has 12 years of schooling right now, i.e., he is a high school graduate. He is deciding whether to invest in a marginal year of schooling to get more skills.
- If he decides not to get the college degree, there is a job that he can work until retirement at wage level, $w_{HS}$. With an extra year of schooling, the job will pay him $w' = (w_{COL} - w_{HS})$ in additional wages until retirement (assume w' is net of loan payments).
- The opportunity cost is $w_{HS}$ dollars in time attending college.

# Human capital model (continued)

- It can be shown that the present discounted value of the additional earnings, w’, is:

$$PV\ (future\ benefits) = w' \frac{1 - e^{-r(T-t)}}{r}$$

*T* is the agent’s retirement date; *t* is the present time, and *r* reflects the agent’s temporal discount rate—with higher *r* meaning less patience. For an agent to be indifferent to a marginal year of schooling, the following must hold:

$$PV\ (future\ benefits) = PV\ (present\ costs) \Leftrightarrow w' \frac{1 - e^{-r(T-t)}}{r} = w_{HS}$$

# Human capital model (continued)

- When *t* is sufficiently far from the retirement date, the second term in the numerator is approximately zero, leading to the following simple relationship between the rate of return to schooling and the individual's temporal discount rate:

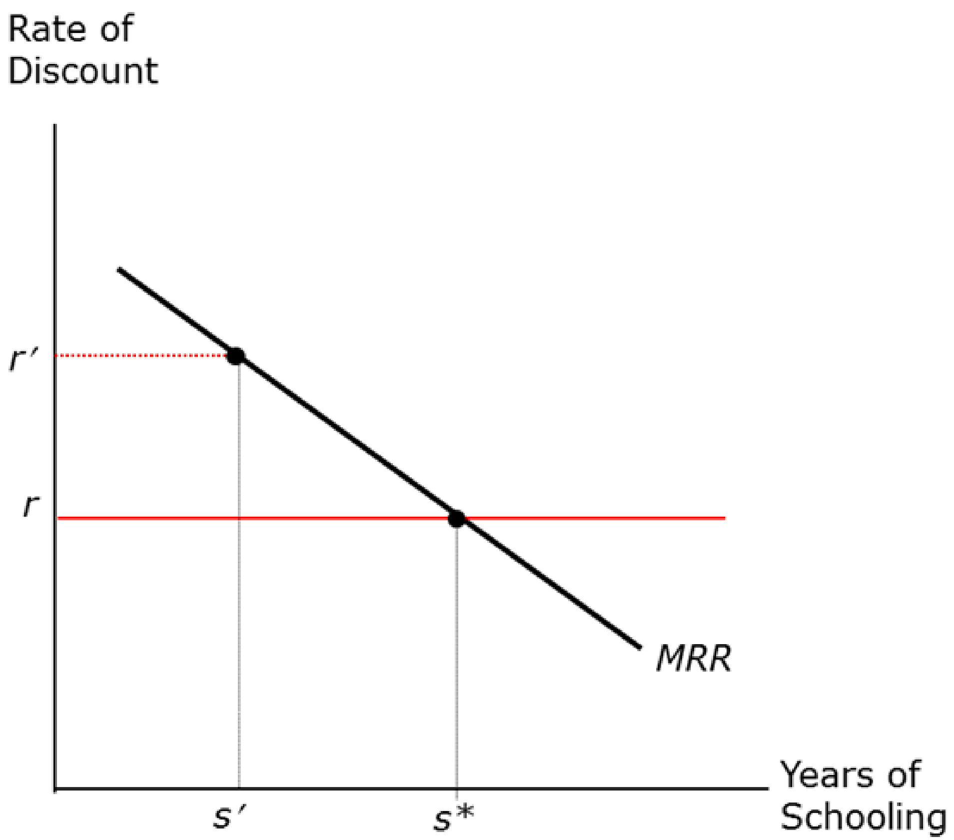
$$w' \frac{1-0}{r} = w_{HS} \rightarrow \frac{w'}{w_{HS}} = r$$

where $\frac{w'}{w_{HS}}$ is the % wage increase from an extra year of schooling.

- Call the % wage premium $\rho$ (Greek letter rho). Note that if rho is greater than *r*, the person should continue schooling because the benefits outweigh the costs. If the reverse is true, it would have been optimal to have one fewer year of schooling.
  - Note also that the smaller *r* is, the smaller is the necessary wage premium to induce the person to stay in school a marginal year.
  - The Borjas textbook calls this "MRR"—the marginal rate of return.

# Diminishing return to schooling

- The rate of return to schooling is positive but diminishing. This means rho is a downward-sloping function of the years of schooling (*s*), i.e., graduating from high school adds more to one's wage than does the first year of college, and so forth.
  - If two people have different levels of patience, $r_1$ and $r_2$, such that $r_1 < r_2$, the more patient (1) will obtain more schooling than the less patient (2).
  - This is like saying that the patient individual has a lower reservation price for investing in human capital; he is willing to take on more years of schooling because he places comparatively large importance on future consumption.



# Rate of return heterogeneity

A marginal year of schooling doesn't have the same return for all individuals.  In fact the entire function $\rho(s)$ can differ among individuals.  Specifically individuals with different initial ability will have a different return functions.

- The rate of return to schooling is also a function of ability, a:  $\rho(s, a)$.
- It is conceivable that higher ability will cause an individual's rho to be <u>either</u> higher or lower than someone with low ability.

$$\frac{\partial \rho}{\partial a_{|s}} \text{ can be greater than or less than zero.}$$

# Ability and return to schooling

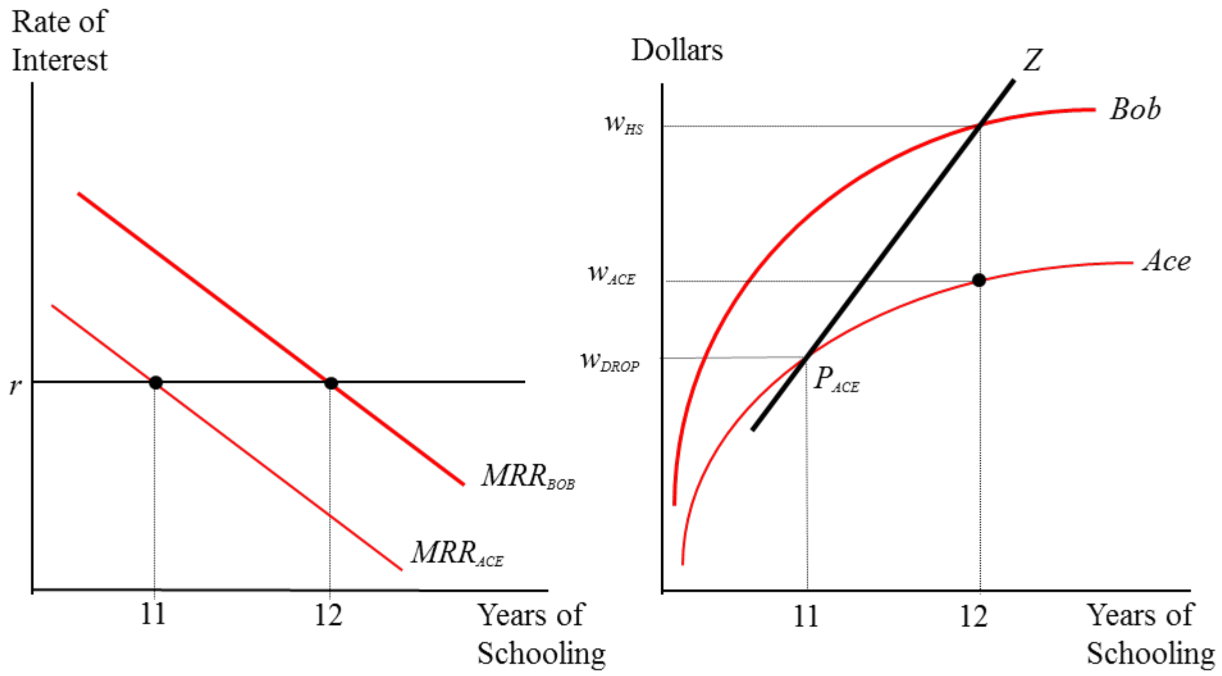
- This is because of two opposing forces.*
  - An individual with higher ability will be more productive—and hence will earn more—than a lower ability individual with the same amount of schooling. So the opportunity cost of additional schooling is larger for high ability individuals.
  - An individual with higher ability will gain more productivity—and wages—from a marginal amount of schooling. The gain from additional schooling is also larger for high ability individuals.
  - Most empirical papers on the subject have concluded that the latter is stronger than the former force—and that high ability individuals have higher rates of return to schooling.**

*More on the theoretical subject of ability and schooling is available (in fact these notes are based on the paper) in Weiss, Yoram (1971). "Ability and the Investment in Schooling: A Theoretical Note on J. Mincer's 'Distribution of Labor Income'." *Journal of Economic Literature*, Vol. 9, No. 2: 459-461.

**Weiss, Andrew (1995). "Human Capital vs. Signaling Explanations for Wages." *Journal of Economic Literature, Vol. 9, No. 4: 133-154*. Particularly pages 137-38.

# Optimal schooling decision

- If ability increases the return to schooling, the rho function for high ability individuals lies above the function for lower ability individuals.
  - Consequently if you consider two individuals with different initial ability and the same rate of time preference, *r*, the individual with higher ability will choose an amount of schooling, $s^*$, that is higher than the $s^*$ for the individual with lower ability.



# Optional: estimating returns to schooling

The mutual relationship among wages, schooling, and ability create one of the most irksome empirical issues in labor economics: ability bias. The optimality condition above, restated in terms of rho:

$$\rho(s, a) = r$$

lends itself especially well to estimating the equilibrium rate of return for a year of schooling. Specifically the logarithmic functional form is implied by this equality:

$$\ln w(s, a) = \ln w(s = 0, a) + rs.$$

The following is analogous to pages 89-94 in the Wooldridge textbook.*

*Wooldridge, Jeffrey M. 2009. *Introductory Econometrics*. 4th edition, South-Western Cengage Learning.

# Estimating using OLS on cross section data

Initially the wage premium for a marginal year of schooling appears easily estimable if one has data on individuals' wage rates and years of schooling. An ordinary least squares (OLS) regression model would look like the following:

$$\ln w_i = \alpha + r s_i + \varepsilon_i$$

Where *i* indexes individuals, *a* and *r* are parameters to be estimated, and epsilon is the (mean zero) error term, which captures individual idiosyncrasies. But one of the assumptions for obtaining unbiased estimates from OLS is that the explanatory variable be uncorrelated with the idiosyncratic error. It is not difficult to show that this assumption is violated if schooling is correlated with ability.

# Ability bias in OLS

- Any variable that is excluded from the regression equation is implicitly included in the error. In this case the error, ε, can be decomposed into two parts.
  - The individual's ability, along with a parameter that measures ability's effect on productivity (and hence, wages).
  - The individual's other idiosyncratic traits that are uncorrelated with ability, schooling, and wages.
  - Formally, consider the possibility that:

$$\varepsilon_i = \theta a_i + \eta_i$$

  where $a_i$ is the ability level of person *i*, $\theta$ (theta) is the rate of return on initial ability, and $\eta$ (eta) is the (now independent) error term, encapsulating idiosyncratic preferences. This model is constructed such that:

$$Covariance(\eta\ , a) = Covariance(\eta\ , s) = 0$$

  - If theories about ability are correct at all, the rate of return on ability will be positive also.

$$\theta > 0$$

# Ability bias (continued)

To see how excluding ability from the regression creates bias, it's useful to think about the predictions that the OLS model generates—namely the predictions of individuals' wages based on their levels of schooling and the estimated parameters, *r* and *a*. These predictions amount to the "best estimate" that can be made of an individual's wage, considering what we know about his schooling. This is called the expectation of wage conditional on schooling. Formally the expectation is expressed as follows:

$$E[\ln w_i | s_i] = \hat{\alpha} + \hat{r} s_i + E[\varepsilon_i | s_i],$$

where the "hats" over *a* and *r* represent the fact that these have been estimated based on data using OLS.

- Were the last term in the conditional expectation zero, we would have no bias in the estimate, and "r hat" would be an unbiased estimate of the effect of a marginal year of schooling on wage (yes!). But if *a* and *s* are (we think positively) correlated, $Covariance(a, s) > 0$; the last term is non-zero.

# Ability bias (continued)

- Observe:

$$E[\varepsilon_i|s_i] = \theta * E[a_i|s_i] + E[\eta_i|s_i]$$

Since eta is uncorrelated with schooling, the last term here is zero. But if schooling and ability are correlated, the first term is non-zero. Specifically it can be shown that the expectation of a conditional on s equals:

$$E[a_i|s_i] = \frac{Covariance\ (a,s)}{Variance\ of\ s} * s_i$$

So,

$$E[\varepsilon_i|s_i] = \theta * \frac{Covariance\ (a,s)}{Variance\ of\ s} * s_i + 0$$

# Bias in estimated return from schooling

- Substituting the above expression into the conditional expectation of wage, we find the bias in the estimate of *r*:

$$E[\ln w_i | s_i] = \hat{\alpha} + \hat{r} s_i + \theta * \frac{Covariance\ (a, s)}{Variance\ of\ s} * s_i$$

and the estimate of the effect of a marginal year of schooling on wage would be:

$$\text{Estimated Causal Effect of Schooling} = \frac{\partial E[\ln w_i | s_i]}{\partial s} = \hat{r} + \theta * \frac{Covariance\ (a, s)}{Variance\ of\ s}$$

- The bias inheres in the fact that this estimate does not converge to the true effect, *r*:

$$\text{Expected Value of Estimated } r = E(\hat{r}) + \theta * \frac{Covariance\ (a, s)}{Variance\ of\ s} = r + bias.$$

# Bias in *r* (continued)

It is also possible to say whether the bias is positive or negative. Since ability and schooling have positive covariance, the bias is positive.

- If we estimate the causal effect of schooling on wage using OLS—and without conditioning on initial ability—the estimated parameter will overestimate the true effect of schooling on wages. Re-arranging the previous line, we get:

$Expected\ Value\ of\ Estimated\ r = r + (bias > 0) \rightarrow\ Expected\ Value\ of\ Estimated\ r > r.$

# Omitted variables problem

Ability bias would not exist if it were possible to observe ability—or something else that is an unbiased estimate of ability. Then you could just condition on schooling and ability to get unbiased estimates of theta and *r*.

$$E[\ln w_i | s_i, a_i] = \hat{\alpha} + \hat{r} s_i + \hat{\theta} a_i + E[\eta_i | s_i, a_i] = \hat{\alpha} + \hat{r} s_i + \hat{\theta} a_i + 0$$

- Ability is inherently difficult to observe. It is not even obvious what the tests (such as IQ and AFQT) that try to measure ability are really measuring.
  - If such a test existed that was an unbiased estimate of productivity-relevant ability, including scores on that test as a control variable in the OLS regression would resolve the bias. This test score would be called a proxy variable.

# Solutions to the omitted variable problem

Absent a proxy variable solution to the ability bias problem, there are two other general methods of circumventing its effects.

- Instrumental Variables (IV) estimation.*
- Using longitudinal data for first difference (FD) or fixed effects (FE) estimation.**

*Wooldridge, chapter 15.

**Wooldridge, chapter 13.

# Instrumental variables (IV)

An instrumental variable is a tool for "purging" the biased variable of interest of its correlation with the unobserved variable.  In order to do so in this case, it must be correlated with schooling but uncorrelated with ability.  The assumption using an IV to resolve ability bias is that z is a valid instrument for schooling if:

$$Covariance(s, z) \neq 0 \; and \; Covariance\ (z, \varepsilon) = 0.$$

- If a suitable IV can be found, estimation can be implemented by the following two step procedure—called Two Stage Least Squares or "2SLS".
  - Regress schooling on *z* and any other control variables used in the wage model.  Generate predicted values of *s* for each observation, i.e.,

    $$E[s_i | z_i, \text{other control variables}] \equiv \hat{s}_i.$$

  - Regress wage (natural log of) on the other control variables and the prediction from the first stage:

    $$\ln w_i = \alpha + r\hat{s}_i + \varepsilon_i$$

# Instrumental variables (continued)

- One of the best-known attempts at this takes advantage of compulsory schooling laws that create age cut-offs for when children have to begin going to school.
  - Children that are on opposite sides of the cut-off should, on average, have identical ability, but the slightly older among them begin school a year earlier.
  - So the calendar month of birth has been used as an instrument  for years of schooling.*

*Angrist, Joshua and Alan Krueger.  1991.  "Does Compulsory Schooling Affect Schooling and Earnings?"  *Quarterly Journal of Economics*: vol. 106:  979-1014.

# Longitudinal or "panel" data

Use of longitudinal data implies observing the same individuals for multiple time periods. It is a powerful tool for addressing omitted variable bias, but less than miraculous when it comes to estimating the causal effect of schooling on wage—for a reason that will become clear. All the variables in the model except for ability (which is person-specific but doesn't change over time) are indexed with a time subscript, $t$, denoting what time period the observation comes from.

$$\ln w_{it} = \alpha + rs_{it} + \theta a_i + \eta_{it}$$

Observations are obtained for two time periods, and the regression model is transformed by subtracting the earlier time period from the later.

$$\Delta \ln w \equiv \ln w_{i1} - \ln w_{i0} = \alpha + rs_{i1} + \theta a_i + \eta_{i1} - (\alpha + rs_{i0} + \theta a_i + \eta_{i0})$$

# Longitudinal data (continued)

- When you perform a transformation like this (first differencing), the temporally invariant terms cancel out.

$$\Delta \ln w = rs_{i1} - rs_{i0} + \eta_{i1} - \eta_{i0} = r(s_{i1} - s_{i0}) + \eta_{i1} - \eta_{i0}$$

The term in parentheses can be characterized as "change in schooling over time" $\Delta s$). If suitable data on wage and schooling over two periods are available, this model will yield unbiased estimates of *r*.

- The problem is that in order to identify the effect of more schooling, <u>you have to actually observe variation in $\Delta s$</u>. This is not common in a data set consisting of adult workers who have completed most of their schooling already.

# Conclusion

Estimating the causal effect of schooling on wages, the wage premium for additional years of schooling, and addressing biases in the estimates comprises one of the largest literatures in economics.  One of the most comprehensive summaries of the various estimates is David Card's "The Causal Effect of Education on Earnings."

- Most estimates congregate in the area of 6 to 10% per year of schooling.

*Card, David.  1999.  in *Handbook of Labor Economics*, Vol. 3:  1801-1863.  Editors:  Orley Ashenfelter and David Card.

# The Logarithmic "Mincerian" Wage Equation

We are interested in showing where the logarithmic functional form for a wage model originates.

$$ln(w_t) = ln(w_0) + \rho t$$

Where *w* is wage rate, the subscripts denote time spent accumulating human capital, e.g., schooling. Rho is the parameter to be estimated, signifying the requisite rate of return on human capital investment to observe an individual who stops at $t$.

The individual is indifferent to continuing and stopping his schooling when the expected present discounted sum of future earnings equals the opportunity cost of acquiring more human capital, i.e., the current wage. Formally,

$$\dot{w}_t \int_t^T e^{-\rho(\tau - t)} d\tau = w_t,$$ where $\dot{w}$ denotes the time derivative of the wage,

i.e., $\dot{w} \equiv \frac{\partial w}{\partial t}$ such that $ln(w_t) - ln(w_0) = \int_0^t \frac{\dot{w}_t}{w_t} d\tau$ and $\frac{d}{dt}(ln\, w_t) = \frac{dln\, w_t}{dw_t}\frac{dw_t}{dt} = \frac{\dot{w}}{w}$

# Wage equation (continued)

Integrating the future wage gains over time gives:

$$\int_t^T e^{-\rho(\tau-t)}d\tau = -\frac{1}{\rho}\left[e^{-\rho(T-t)} - e^{-\rho(t-t)}\right] = \frac{1}{\rho}\left[1 - e^{-\rho(T-t)}\right]$$

So the equilibrium condition is,

$$\dot{w}_t\frac{1}{\rho}\left[1 - e^{-\rho(T-t)}\right] = w_t \Leftrightarrow \frac{\dot{w}_t}{w_t} = \frac{\rho}{1 - e^{-\rho(T-t)}}$$

This is the line at the top of the 8th slide, showing the present discounted returns to schooling.

If the future time horizon is large, i.e., T is far from t,

$$\frac{\dot{w}_t}{w_t} = \frac{\rho}{1-0} = \rho$$

# Wage equation (concluded)

Use the property of integrating the percentage changes over time to arrive at the wage equation. Integrating both sides gives the following. Rewrite with log of current wage on the left side.

$$\int_0^t \frac{\dot{w}_t}{w_t} dt = \int_0^t \rho \, dt \Leftrightarrow \ln(w_t) - \ln(w_0) = \rho t \Leftrightarrow \ln(w_t) = \ln(w_0) + \rho t$$

Back.