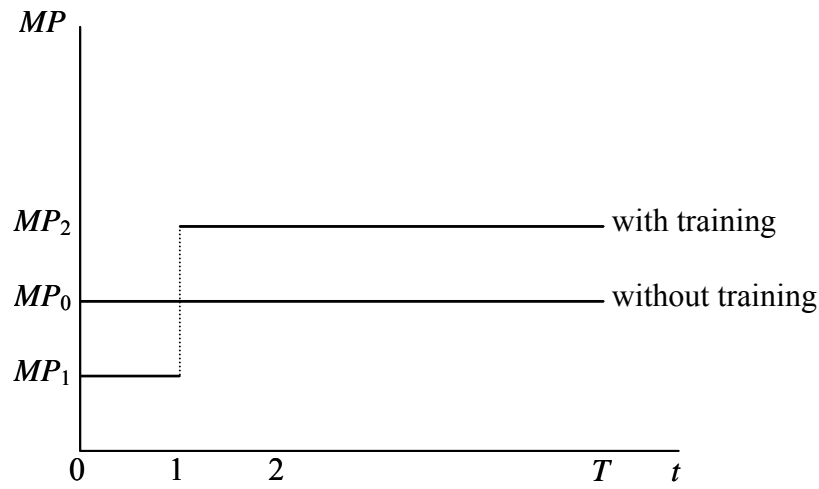


Human Capital – Training

- A. General model of post-school skills investments
 - 1. A substantial amount of skills acquisition occurs after people leave school
 - a. some of this involves “learning by doing” – proficiency at job tasks increases simply by performing the tasks over and over; does not require any special action by the employer (we are not going to consider this)
 - b. some jobs also require or provide additional
 - 1) classroom training – more generally, training away from the job task
 - 2) on-the-job training – training that occurs while performing a job task
 - c. we will examine an employer’s decisions whether to provide these types of training
 - 2. Assumptions
 - a. will consider discrete time periods
 - b. job tenure is finite, lasts T periods
 - c. training will take one period and will occur at the beginning of an employee’s job tenure
 - d. no direct cost of training
 - e. productivity of an untrained employee is constant at MP_0 per period
 - f. productivity of a trained employee is:
 - 1) $MP_1 < MP_0$ during training
 - 2) $MP_2 > MP_0$ in each period after training occurs

g. graph of the marginal productivities



h. present value of training

1) for an untrained employee

$$PV_0 = MP_0 + \sum_{t=2}^T (1+r)^{1-t} MP_0$$

2) for a trained employee

$$PV_{TR} = MP_1 + \sum_{t=2}^T (1+r)^{1-t} MP_2$$

3. Training is worthwhile to the firm if $PV_{TR} > PV_0$ or

$$\sum_{t=2}^T (1+r)^{1-t} (MP_2 - MP_0) > MP_0 - MP_1$$

term on the left is the present discounted value of the marginal increment to productivity (marginal benefit);
term on the right is the short-term loss in productivity associated with training (opportunity cost)

4. How much will the firm pay the employee in each period?

- a. assuming competitive markets, employer faces a profitability constraint

$$\sum_{t=1}^T (1+r)^{1-t} W_t \leq \sum_{t=1}^T (1+r)^{1-t} MP_t$$

- b. constraint is clearly met if employer pays employee exactly her marginal product
- c. when can the employer pay $W_1 > MP_1$ in the first period—that is, subsidize the cost of training? can only do this if the employer can pay $W_2 < MP_2$ in later periods
- d. this second condition cannot occur if the worker's post-training productivity to other firms is MP_2 ; if this were the case, the worker could leave the firm providing the training for another firm and receive wages equal to MP_2
5. Lesson is that *the type of training matters*
- a. pure general training – training that increases productivity at all firms (e.g., like learning a computer programming language or understanding general principles of accounting)
- b. pure specific training – training that increases productivity only at a single firm (e.g., learning a company's special software packages or its internal accounting practices)
- c. for pure general training, the employee bears all of the cost of training
- d. for pure specific training, the firm is in a position to absorb the costs of training through a flatter wage profile

- 1) so long as the present discounted value of the wage stream for trained workers is higher than that for untrained workers, workers will have an incentive to undergo training
 - 2) so long as the wages at the current firm equal or exceed those offered by other firms, trained workers stay
 - 3) note that the firm is in a position to capture the benefits associated with training
- e. investments and returns associated with specific training may partially explain why firms try to hold onto employees during economic downturns and why firms use lay-off policies rather than outright dismissals during downturns

B. Empirical evidence regarding employer-provided training

1. Data requirements to examine these issues directly
 - a. need indicators of training (whether and how it occurred, how long it took)
 - b. need indicators of the direct costs
 - c. need measures of wages before, during and after training
 - d. information about the training itself is best provided by firms
 - e. longitudinal information about wages is usually provided by workers
 - f. hard to obtain all of this information together; as a result, relatively few studies have looked at direct evidence

2. Most research is based on indirect evidence, usually analyses of how wages vary with job tenure and experience
 - a. in a seminal paper, Mincer (1962) calculated
 - 1) the value of training by examining the amount of experience needed for the earnings of people with one level of education to catch up with the earnings of people with more education
 - 2) returns using data on apprenticeships and data on physician's training in medical specialties
 - b. Brown (1989) linked wage-tenure profiles with characteristics (average training requirements) of professions; found most wage growth occurred during "training periods"
 - c. Topel (1991)
 - 1) estimated fixed effects models of wage growth that accounted for (a) person-specific effects (estimated from multiple wage observations, including observations across firms) and (b) match-specific effects (estimated from multiple wage observations within a firm)
 - 2) found that most wage growth occurs early in peoples' job tenures – 10 years of job seniority worth about 25 percent more in wages
3. Problems with studies based on indirect evidence
 - a. problems with omitted variables could lead to spurious profiles (e.g., people with lower productivity and lower wages more likely to leave over time leading to longer-tenure workers being selectively more productive and higher paid)

- b. other explanations for upward-sloping wage-tenure profiles
 - 1) back-loaded compensation to encourage workers to stay with firms
 - 2) schemes to get mobile workers to reveal themselves

C. Public training programs (taken from Heckman, LaLonde and Smith 1999)

1. Characteristics

- a. important component of social policy – expenditures vary across countries; however, in many large OECD countries expenditures for training are nearly as large as expenditures on direct income support (e.g., unemployment insurance payments)
- b. intended to increase skills among people who are “structurally” disadvantaged because of either a lack of skills or a change in the demand for their skills
- c. encompasses many types of programs
 - 1) classroom training (both basic skills and vocational skills)
 - 2) subsidized wage programs and subsidized on-the-job training
 - 3) job search assistance
- d. lots of heterogeneity in other aspects
 - 1) different levels of intensity – in U.S., training programs typically run for 3-4 months; however, in some other countries, programs run as short as a few weeks

- 2) different settings – in U.S., programs are sometimes run by local agencies, subcontracted to vocational schools, or provided through community colleges
 - 3) different types of focus – in U.S., many programs are run by local councils that include business, labor and social service representatives; in other countries a more standardized curriculum is used
 - 4) different types of clients
 - e. heterogeneity in program features almost certainly implies that there will be heterogeneity in outcomes
2. evaluation problem
- a. let D be a binary variable which indicates “treatment” – this could mean receiving some service vs. no service but could also mean just receiving different services
 - b. potential outcomes for each person
 - 1) Y_0 – if untreated ($D = 0$)
 - 2) Y_1 – if treated ($D = 1$)
 - c. let X represent a set of conditioning variables
 - d. let
 - 1) $Y_0 = \mu_0(X) + U_0$
 - 2) $Y_1 = \mu_1(X) + U_1$
 - 3) where $\mu_j(X) = E(Y_j | X)$ and $E(U_j | X) = 0$ for $j=0,1$ (second assumption made for convenience)
 - 4) also assume that D does not determine X given the potential outcomes

- e. gain from treatment is $\Delta = Y_1 - Y_0$
 - f. problem is that this is not observable because person can only be observed in one state at a time
 - g. to characterize Δ , we need to come up with counterfactuals; however, there are many types of counterfactual situations
3. some counterfactuals
- a. use linear switching regression framework in which
 - 1) observed outcome $Y = DY_1 - (1-D)Y_0$
 - 2) $\mu_j(X) = X\beta_j$
 - b. can show

$$Y = X\beta_0 + DX(\beta_1 - \beta_0) + \{U_0 + D(U_1 - U_0)\}$$
 - c. “regression” model with X and interactions of D and X as observed controls; attempts to use aggregate data to measure the average direct impact of treatment
 - d. in restricted (and unrealistic) case where all slope coefficients are equal across the two regimes, model simplifies to

$$Y = X\beta_0 + D\alpha + \{U_0 + D(U_1 - U_0)\}$$
 - e. possible problems running this regression because

$$E[U_0 + D(U_1 - U_0)] = E(U_1 - U_0 | D = 1) \Pr(D = 1)$$
 but $E(U_1 - U_0 | D = 1)$ may not equal zero – for instance, if participation in the treatment depends on the expected gains
 - f. in practice, most studies estimate a different impact, the impact of treatment on the treated or

$$E(\Delta | X, D = 1)$$

g. in the linear switching model this is equivalent to

$$X(\beta_1 - \beta_0) + E[U_1 - U_0 | X, D = 1]$$

which is only the same as the general treatment effect if $E[U_1 - U_0 | X, D = 1] = 0$ (e.g., agents don't act on gains from participating in a program)

4. some general types of estimation approaches

a. assume $E[U_1 - U_0 | X, D = 1] = 0$, this is done when people estimate standard regression equations

b. use before and after comparisons

1) assumes mean outcome in the no-treatment state (conditional on the X s) is the same before and after treatment

2) requires before and after data

3) can run into problems such as a pre-treatment dip in the outcomes

4) with longitudinal data, essentially use a person's earlier outcomes as a control for his/her later outcomes

c. matching estimators

1) construct control groups based on observably similar individuals

2) requires special methods when there are lots of observable characteristics

3) need to be able to find matches across a range of X s

4) does not control for differences in unobserved characteristics

5) does, however, force researcher to be explicit about the basis for comparisons

- d. control function estimators
 - 1) rely on assumptions regarding the conditional distribution of the unobservables
 - 2) address selection based on unobservables
 - 3) examples are the two-stage and MLE switching regression models
- e. instrumental variable estimators
 - 1) also addresses selection based on the unobservables
 - 2) of course, need a valid instrument for D
 - 3) can be difficult to apply if the impacts of treatment vary across people
- f. random-assignment experiments
 - 1) essentially balances selection bias across people who are randomly assigned into and out of treatment
 - 2) important when randomization occurs – eligibility, application or acceptance
 - 3) relatively expensive
 - 4) may be hard to gain participation from program officials
 - 5) scale issues
 - 6) may be hard to extrapolate from experiments
- 5. lessons from Heckman, LaLonde and Smith
 - a. many outcomes of interest that can be evaluated – appropriate estimator depends on the outcome of interest

- b. no preferred method – methods depend on the program, outcome of interest, theory and data
- c. good data help in evaluating outcomes
- d. need valid comparison groups
- e. different estimators are sensitive to different underlying problems in the data and different sources of selection
- f. need to be sensitive to general equilibrium effects

References

- Becker, Gary. *Human Capital* 3rd edition. Chicago: University of Chicago Press, 1993.
- Brown, James. “Why Do Wages Increase with Tenure? On-the-Job Training and Life-Cycle Wage Growth Observed within Firms.” *American Economic Review* 79:5 (December 1989), 971-91.
- Heckman, James, Robert LaLonde and Jeffrey Smith. “The Economics and Econometrics of Active Market Labor Programs.” In *Handbook of Labor Economics* vol. 3A, edited by O. Ashenfelter and D. Card. Amsterdam: North-Holland, 1999.
- Mincer, Jacob. “On-the-Job Training: Costs, Returns, and Some Implications.” *Journal of Political Economy* 70:5 part 2 (October 1962), 50-79.
- Topel, Robert. “Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority.” *Journal of Political Economy* 99:1 (February 1991), 145-76.