Volume Title: Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education, First Edition

Volume Author/Editor: Gary S. Becker

Volume Publisher: NBER

Volume ISBNs:

Volume URL: http://www.nber.org/books/beck-5

Conference Date: n/a

Publication Date: 1964

Chapter Title: Investment in Human Capital: Effects on Earnings

Chapter Author(s): Gary S. Becker

Chapter URL: http://www.nber.org/chapters/c14408

Chapter pages in book: (p. 7 – 36)

CHAPTER II

Investment in Human Capital: Effects on Earnings[1]

THE original aim of this study was to estimate the money rate of return to college and high-school education in the United States. In order to set these estimates in the proper context, a brief formulation of the theory of investment in human capital was undertaken. It soon became clear to me, however, that more than a restatement was called for; while important and pioneering work had been done on the economic return to various occupations and education classes,[2] there had been few, if any, attempts to treat the process of investing in people from a general viewpoint or to work out a broad set of empirical implications. I began then to prepare a general analysis of investment in human capital.

It eventually became apparent that this general analysis would do much more than fill a gap in formal economic theory: it offers a unified explanation of a wide range of empirical phenomena which have either been given *ad hoc* interpretations or have baffled investigators. Among these phenomena are the following: (1) Earnings typically increase with age at a decreasing rate. Both the rate of increase and

[1] This chapter and the one that follows were published in somewhat different form in *Investment in Human Beings,* NBER Special Conference 15, supplement to *Journal of Political Economy,* October 1962, pp. 9-49.

[2] In addition to the earlier works of Smith, Mill, and Marshall, see the brilliant work (which greatly influenced my own thinking about occupational choice) by M. Friedman and S. Kuznets, *Income from Independent Professional Practice,* New York, NBER, 1945; see also H. Clark, *Life Earnings in Selected Occupations in the U.S.,* New York, 1937; J. R. Walsh, "Capital Concept Applied to Man," *Quarterly Journal of Economics,* February 1935; G. Stigler and D. Blank, *The Demand and Supply of Scientific Personnel,* New York, NBER, 1957. In recent years, of course, there has been considerable work, especially by T. W. Schultz; see, for example, his "Investment in Human Capital," *American Economic Review,* March 1961, pp. 1-17.

the rate of retardation tend to be positively related to the level of skill. (2) Unemployment rates tend to be inversely related to the level of skill. (3) Firms in underdeveloped countries appear to be more "paternalistic" toward employees than those in developed countries. (4) Younger persons change jobs more frequently and receive more schooling and on-the-job training than older persons do. (5) The distribution of earnings is positively skewed, especially among professional and other skilled workers. (6) Abler persons receive more education and other kinds of training than others. (7) The division of labor is limited by the extent of the market. (8) The typical investor in human capital is more impetuous and thus more likely to err than is the typical investor in tangible capital.

What a diverse and even confusing array! Yet all these, as well as many other important empirical implications, can be derived from very simple theoretical arguments. The purpose here is to set out these arguments in general form, with the emphasis placed on empirical implications, although little empirical material is presented. Systematic empirical work appears in Part Two.

In this chapter a lengthy discussion of on-the-job training is presented and then, much more briefly, discussions of investment in schooling, information, and health. On-the-job training is dealt with so elaborately not because it is more important than other kinds of investment in human capital—although its importance is often underrated—but because it clearly illustrates the effect of human capital on earnings, employment, and other economic variables. For example, the close connection between indirect and direct costs and the effect of human capital on earnings at different ages are vividly brought out. The extended discussion of on-the-job training paves the way for much briefer discussions of other kinds of investment in human beings.

1. On-the-Job Training

Theories of firm behavior, no matter how they differ in other respects, almost invariably ignore the effect of the productive process itself on worker productivity. This is not to say that no one recognizes that productivity is affected by the job itself; but the recognition has not been formalized, incorporated into economic analysis, and its implications worked out. I now intend to do just that, placing special emphasis on the broader economic implications.

Many workers increase their productivity by learning new skills and perfecting old ones while on the job. Presumably, future productivity can be improved only at a cost, for otherwise there would be an unlimited demand for training. Included in cost are the value placed on the time and effort of trainees, the "teaching" provided by others, and the equipment and materials used. These are costs in the sense that they could have been used in producing current output if they had not been used in raising future output. The amount spent and the duration of the training period depend partly on the type of training since more is spent for a longer time on, say, an intern than a machine operator.

Consider explicitly now a firm that is hiring employees for a specified time period (in the limiting case this period approaches zero), and for the moment assume that both labor and product markets are perfectly competitive. If there were no on-the-job training, wage rates would be given to the firm and would be independent of its actions. A profit-maximizing firm would be in equilibrium when marginal products equaled wages, that is, when marginal receipts equaled marginal expenditures. In symbols

$$MP = W, \tag{1}$$

where W equals wages or expenditures and MP equals the marginal product or receipts. Firms would not worry too much about the relation between labor conditions in the present and future, partly because workers would only be hired for one period and partly because wages and marginal products in future periods would be independent of a firm's current behavior. It can therefore legitimately be assumed that workers have unique marginal products (for given amounts of other inputs) and wages in each period, which are, respectively, the maximum productivity in all possible uses and the market wage rate. A more complete set of equilibrium conditions would be the set

$$MP_t = W_t, \tag{2}$$

where t refers to the tth period. The equilibrium position for each period would depend only on the flows during that period.

These conditions are altered when account is taken of on-the-job training and the connection thereby created between present and future receipts and expenditures. Training might lower current re-

ceipts and raise current expenditures, yet firms could profitably provide this training if future receipts were sufficiently raised or future expenditures sufficiently lowered. Expenditures during each period need not equal wages, receipts need not equal the maximum possible marginal productivity, and expenditures and receipts during all periods would be interrelated. The set of equilibrium conditions summarized in equation (2) would be replaced by an equality between the *present values* of receipts and expenditures. If E_t and R_t represent expenditures and receipts during period t, and i the market discount rate, then the equilibrium condition can be written as

$$\sum_{t=0}^{n-1} \frac{R_t}{(1+i)^{t+1}} = \sum_{t=0}^{n-1} \frac{E_t}{(1+i)^{t+1}}, \tag{3}$$

when n represents the number of periods, and R_t and E_t depend on all other receipts and expenditures. The equilibrium condition of equation (2) has been generalized, for if marginal product equals wages in each period, the present value of the marginal product stream would have to equal the present value of the wage stream. Obviously, however, the converse need not hold.

If training were given only during the initial period, expenditures during the initial period would equal wages plus the outlay on training, expenditures during other periods would equal wages alone, and receipts during all periods would equal marginal products. Equation (3) becomes

$$MP_0 + \sum_{t=1}^{n-1} \frac{MP_t}{(1+i)^t} = W_0 + k + \sum_{t=1}^{n-1} \frac{W_t}{(1+i)^t}, \tag{4}$$

where k measures the outlay on training.

If a new term is defined,

$$G = \sum_{t=1}^{n-1} \frac{MP_t - W_t}{(1+i)^t}, \tag{5}$$

equation (4) can be written as

$$MP_0 + G = W_0 + k. \tag{6}$$

Since the term k only measures the actual outlay on training, it does not entirely measure training costs, for it excludes the time that a person spends on this training, time that could have been used to produce current output. The difference between what could have been produced, MP_0', and what is produced, MP_0, is the opportunity cost of the time spent in training. If C is defined as the sum of opportunity costs and outlays on training, (6) becomes

$$MP_0' + G = W_0 + C. \tag{7}$$

The term G, the excess of future receipts over future outlays, is a measure of the return to the firm from providing training; and, therefore, the difference between G and C measures the difference between the return from and the cost of training. Equation (7) shows that the marginal product would equal wages in the initial period only when the return equals costs, or G equals C; it would be greater or less than wages as the return was smaller or greater than costs. Those familiar with capital theory might argue that this generalization of the simple equality between marginal product and wages is spurious because a full equilibrium would require equality between the return from an investment—in this case, made on the job—and costs. If this implied that G equals C, marginal product would equal wages in the initial period. There is much to be said for the relevance of a condition equating the return from an investment with costs, but such a condition does not imply that G equals C or that marginal product equals wages. The following discussion demonstrates that great care is required in the application of this condition to on-the-job investment.

Our treatment of on-the-job training produced some general results —summarized in equations (3) and (7)—of wide applicability, but more concrete results require more specific assumptions. In the following sections two types of on-the-job training are discussed in turn: general and specific.

GENERAL TRAINING

General training is useful in many firms besides those providing it; for example, a machinist trained in the army finds his skills of value in steel and aircraft firms, and a doctor trained (interned) at one hospital finds his skills useful at other hospitals. Most on-the-job training presumably increases the future marginal productivity of

workers in the firms providing it; general training, however, also increases their marginal product in many other firms as well. Since in a competitive labor market the wage rates paid by any firm are determined by marginal productivities in other firms, future wage rates as well as marginal products would increase in firms providing general training. These firms could capture some of the return from training only if their marginal product rose by more than their wages. "Perfectly general" training would be equally useful in many firms and marginal products would rise by the same extent in all of them. Consequently, wage rates would rise by exactly the same amount as the marginal product and the firms providing such training could not capture any of the return.

Why, then, would rational firms in competitive labor markets provide general training if it did not bring any return? The answer is that firms would provide general training only if they did not have to pay any of the costs. Persons receiving general training would be willing to pay these costs since training raises their future wages. Hence it is the trainees, not the firms, who would bear the cost of general training and profit from the return.[3]

These and other implications of general training can be more formally demonstrated in equation (7). Since wages and marginal products are raised by the same amount, MP_t must equal W_t for all $t = 1, \ldots n-1$, and therefore

$$G = \sum_{t=1}^{n-1} \frac{MP_t - W_t}{(1+i)^t} = 0. \tag{8}$$

Equation (7) is reduced to

$$MP_0' = W_0 + C, \tag{9}$$

or

$$W_0 = MP_0' - C. \tag{10}$$

[3] Some persons have asked why any general training is provided if firms do not collect any of the returns. The answer is simply that they have an incentive to do so wherever the demand price for training is at least as great as the supply price or cost of providing the training. Workers in turn would prefer to be trained on the job rather than in specialized firms (schools) if the training and work complemented each other (see the discussion in section 2 below).

In terms of actual marginal product

$$MP_0 = W_0 + k, \tag{9'}$$

or

$$W_0 = MP_0 - k. \tag{10'}$$

The wage of trainees would not equal their opportunity marginal product but would be less by the total cost of training. In other words, employees would pay for general training by receiving wages below their current (opportunity) productivity. Equation (10) has many other implications, and the rest of this section is devoted to developing the more important ones.

Some might argue that a really "net" definition of marginal product, obtained by subtracting training costs from "gross" marginal product, must equal wages even for trainees. Such an interpretation of net productivity could formally save the equality between marginal product and wages here, but not always, as shown later. Moreover, regardless of which interpretation is used, training costs would have to be included in any study of the relation between wages and productivity.

Employees pay for general on-the-job training by receiving wages below what they could receive elsewhere. "Earnings" during the training period would be the difference between an income or flow term (potential marginal product) and a capital or stock term (training costs), so that the capital and income accounts would be closely intermixed, with changes in either affecting wages. In other words, earnings of persons receiving on-the-job training would be net of investment costs and would correspond to the definition of *net* earnings used throughout this paper, which subtracts all investment costs from "gross" earnings. Therefore, our departure with this definition of earnings from the accounting conventions used for transactions in material goods—which separate income from capital accounts to prevent a transaction in capital from *ipso facto*[4] affecting the income side—is not capricious but is grounded in a fundamental difference between the way investment in material and human capital are "written off." The underlying cause of this difference undoubtedly is the widespread reluctance to treat people as capital and the accompanying tendency to treat all wage receipts as earnings.

[4] Of course, a shift between assets with different productivities would affect the income account on material goods even with current accounting practices.

Intermixing the capital and income accounts could make the reported "incomes" of trainees unusually low and perhaps negative, even though their long-run or lifetime incomes were well above average. Since a considerable fraction of young persons receive some training, and since trainees tend to have lower current and higher subsequent earnings than other youth, the correlation of current consumption with the current earnings of young males[5] would not only be much weaker than the correlation with long-run earnings, but the signs of these correlations might even differ.[6]

Doubt has been cast on the frequent assertion that no allowance is made in the income accounts for depreciation on human capital.[7] A depreciation-type item is deducted, at least from the earnings due to on-the-job training, for the cost would be deducted during the training period. Depreciation on tangible capital does not bulk so large in any one period because it is usually "written off" or depreciated during a period of time designed to approximate its economic life. Hence human and tangible capital appear to differ more in the time pattern of depreciation than in its existence,[8] and the effect on wage income of a rapid "write-off" of human capital is what should be emphasized and studied.

This point can be demonstrated differently and more rigorously. The ideal depreciation on a capital asset during any period would equal its change in value during the period. In particular, if value

[5] The term "young males" rather than "young families" is used because as J. Mincer has shown (in his "Labor Force Participation of Married Women," *Aspects of Labor Economics*, Princeton for NBER, 1962), the labor force participation of wives is positively correlated with the difference between a husband's long-run and current income. Participation of wives, therefore, makes the correlation between a family's current and a husband's long-run income greater than that between a husband's current and long-run income.

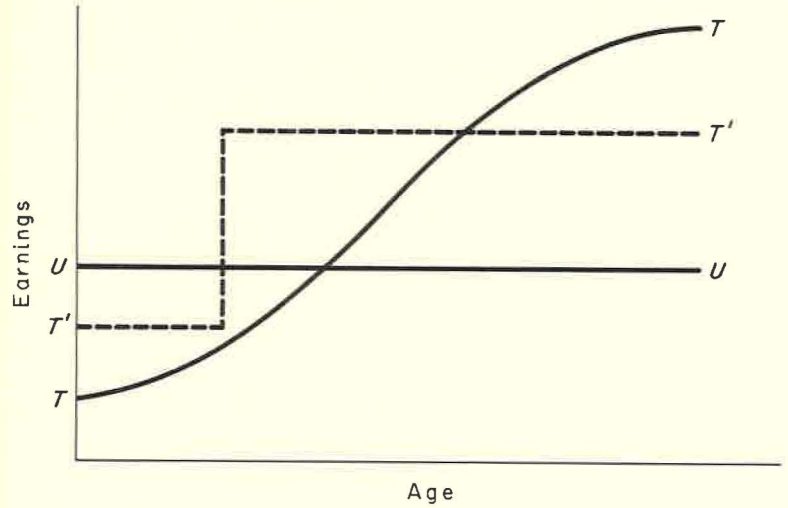
[6] A difference in signs is impossible in Friedman's analysis of consumer behavior because he assumes that, at least in the aggregate, transitory and long-run (that is, permanent) incomes are uncorrelated (see his *A Theory of the Consumption Function*, Princeton for NBER, 1957); I am suggesting that they may be *negatively* correlated for young persons.

[7] See C. Christ, "Patinkin on Money, Interest, and Prices," *Journal of Political Economy*, August 1957, p. 352; and W. Hamburger, "The Relation of Consumption to Wealth and the Wage Rate," *Econometrica*, January 1955.

[8] In a recent paper, R. Goode has argued (see his "Educational Expenditures and the Income Tax," in Selma J. Mushkin, ed., *Economics of Higher Education*, Washington, 1962) that educated persons should be permitted to subtract from income a depreciation allowance on tuition payments. Such an allowance is apparently not required for on-the-job training costs or, as seen later, for the indirect costs of education; indeed, one might argue, on the contrary, that too much or too rapid depreciation is permitted on such investments.

CHART 1

Relation of Earnings to Age



rose, a negative depreciation term would have to be subtracted or a positive appreciation term added to the income from the asset. Since training costs would be deducted from earnings during the training period, the economic "value" of a trainee would at first increase rather than decrease with age, and only later begin to decrease. Therefore, a negative rather than a positive depreciation term would have to be subtracted initially.[9]

Training has an important effect on the relation between earnings and age. Suppose that untrained persons received the same earnings regardless of age, as shown by the horizontal line UU in Chart 1. Trained persons would receive lower earnings during the training period because training is paid for at that time, and higher earnings at later ages because the return is collected then. The combined effect of paying for and collecting the return from training in this way would be to make the age earnings curve of trained persons, shown by TT in Chart 1, steeper than that of untrained persons, the difference being greater the greater the cost of, and return from, the investment.

[9] See Chapter VII, section 2, for some empirical estimates of "depreciation" on human capital.

Not only does training make the curve steeper but, as indicated by Chart 1, also more concave; that is, the rate of increase in earnings is affected more at younger than at older ages. Suppose, to take an extreme case, that training raised the level of marginal productivity but had no effect on the slope, so that the marginal productivity of trained persons was also independent of age. If earnings equaled marginal product, TT would merely be parallel to and higher than UU, showing neither slope nor concavity. Since, however, earnings of trained persons would be below marginal productivity during the training period and equal afterward, they would rise sharply at the end of the training period and then level off (as shown by the dashed line $T'T'$ in Chart 1), imparting a concave appearance to the curve as a whole. In this extreme case an extreme concavity appears (as in TT); in less extreme cases the principle would be the same and the concavity more continuous.

Foregone earnings are an important, although neglected, cost of much investment in human capital and should be treated in the same way as direct outlays. Indeed, *all* costs appear as foregone earnings to workers receiving on-the-job training; that is, all costs appear as lower earnings than could be received elsewhere, although direct outlays, C, may really be an important part of costs. The arbitrariness of the division between indirect and direct costs and the resulting advantage of treating total costs as a whole[10] can be further demonstrated by contrasting school and on-the-job training. Usually only the direct costs of school training are emphasized, even though opportunity costs are sometimes (as with college education) an important part of the total. A shift from school training to on-the-job training would, however, reverse the emphasis and make all costs appear as foregone earnings, even when direct outlays were important.

Income-maximizing firms in competitive labor markets would not pay the cost of general training and would pay trained persons the market wage. If, however, training costs were paid, many persons would seek training, few would quit during the training period, and labor costs would be relatively high. Firms that did not pay trained persons the market wage would have difficulty satisfying their skill

[10] The equivalence between indirect and direct costs applies to consumption as well as to investment decisions. In my paper, *A Theory of the Allocation of Time*, IBM Research Paper RC 1149, March 20, 1964, an analysis incorporating both direct and indirect consumption costs is applied to the choice between work and nonwork, price and income elasticities of demand for goods, the economic function of queues, and several other areas.

requirements and would also tend to be less profitable than other firms. Firms that paid both for training and less than the market wage for trained persons would have the worst of both worlds, for they would attract too many trainees and too few trained persons.

These principles have been clearly demonstrated during the last few years in discussions of problems in recruiting military personnel. The military offers training in a wide variety of skills and many are very useful in the civilian sector. Training is provided during part or all of the first enlistment period and used during the remainder of the first period and hopefully during subsequent periods. This hope, however, is thwarted by the fact that re-enlistment rates tend to be inversely related to the amount of civilian-type skills provided by the military.[11] Persons with these skills leave the military more readily because they can receive much higher wages in the civilian sector. Net military wages for those receiving training are higher relative to civilian wages during the first than during subsequent enlistment periods because training costs are largely paid by the military. Not surprisingly, therefore, first-term enlistments for skilled jobs are obtained much more easily than are re-enlistments.

The military is a conspicuous example of an organization that both pays at least part of training costs and does not pay market wages to skilled personnel. It has had, in consequence, relatively easy access to "students" and heavy losses of "graduates." Indeed, its graduates make up the predominant part of the supply in several civilian occupations. For example, well over 90 per cent of United States commercial airline pilots received much of their training in the armed forces. The military, of course, is not a commercial organization judged by profits and losses and has had no difficulty surviving and even thriving.

What about the old argument that firms in competitive labor markets have no incentive to provide on-the-job training because trained workers would be bid away by other firms? Firms that train workers are supposed to impart external economies to other firms because the latter can use these workers free of any training charge. An analogy with research and development is often drawn since a firm developing a process that cannot be patented or kept secret would impart ex-

[11] See *Manpower Management and Compensation*, report of the Cordiner Committee, Washington, 1957, Vol. I, Chart 3, and the accompanying discussion. The military not only wants to eliminate the inverse relation but apparently would like to create a positive relation because they have such a large investment in heavily trained personnel. For a recent and excellent study, see Gorman C. Smith, "Differential Pay for Military Technicians," unpublished Ph.D. dissertation, Columbia University, 1964.

ternal economies to competitors. This argument and analogy would apply if firms were to pay training costs, for they would suffer a "capital loss" whenever trained workers were bid away by other firms. Firms can, however, shift training costs to trainees and have an incentive to do so when faced with competition for their services.[12]

The difference between investment in training and in research and development can be put very simply. Without patents or secrecy, firms in competitive industries cannot establish property rights in innovations, and these innovations become fair game for all comers. Patent systems try to establish these rights so that incentives can be provided to invest in research. Property rights in skills, on the other hand, are automatically vested, for a skill cannot be used without permission of the person possessing it. The property right of the worker in his skills is the source of his incentive to invest in training by accepting a reduced wage during the training period and explains why an analogy with unowned innovations is misleading.

SPECIFIC TRAINING

Completely general training increases the marginal productivity of trainees by exactly the same amount in the firms providing the training as in other firms. Clearly some kinds of training increase productivity by different amounts in the firms providing the training and in other firms. Training that increases productivity more in firms providing it will be called specific training. Completely specific training can be defined as training that has no effect on the productivity of trainees that would be useful in other firms. Much on-the-job training is neither completely specific nor completely general but increases productivity more in the firms providing it and falls within the definition of specific training. The rest increases productivity by at least as much in other firms and falls within a definition of general training. A few

[12] Sometimes the alleged external economies from on-the-job training have been considered part of the "infant industry" argument for protection (see J. Black "Arguments for Tariffs," *Oxford Economic Papers,* June 1959, pp. 205-206). Our analysis suggests, however, that the trouble tariffs are supposed to overcome must be traced back to difficulties that workers have in financing investment in themselves—in other words, to ignorance or capital market limitations that apply to expenditures on education, health, as well as on-the-job training. Protection would serve the same purpose as the creation of monopsonies domestically, namely, to convert general into specific capital so that firms can be given an incentive to pay for training (see the remarks on specific training below and in section 4 of this chapter). Presumably a much more efficient solution would be to improve the capital market directly through insurance of loans, subsidies, information, etc.

illustrations of the scope of specific training are presented before a formal analysis is developed.

The military offers some forms of training that are extremely useful in the civilian sector, as already noted, and others that are only of minor use to civilians, i.e., astronauts, fighter pilots, and missile men. Such training falls within the scope of specific training because productivity is raised in the military but not (much) elsewhere.

Resources are usually spent by firms in familiarizing new employees with their organization,[13] and the knowledge thus acquired is a form of specific training because productivity is raised more in the firms acquiring the knowledge than in other firms. Other kinds of hiring costs, such as employment agency fees, the expenses incurred by new employees in finding jobs, or the time employed in interviewing, testing, checking references, and in bookkeeping do not so obviously increase the knowledge of new employees, but they too are a form of specific investment in human capital, although not training. They are an investment because outlays over a short period create distributed effects on productivity; they are specific because productivity is raised primarily in the firms making the outlays; they are in human capital because they lose their value whenever employees leave. In the rest of this section reference is mostly to on-the-job specific training even though the analysis applies to all on-the-job specific investment.

Even after hiring costs are incurred, firms usually know only a limited amount about the ability and potential of new employees. They try to increase their knowledge in various ways—testing, rotation among departments, trial and error, etc.—for greater knowledge permits a more efficient utilization of manpower. Expenditures on acquiring knowledge of employee talents would be a specific investment if the knowledge could be kept from other firms, for then productivity would be raised more in the firms making the expenditures than elsewhere.

The effect of investment in employees on their productivity elsewhere depends on market conditions as well as on the nature of the investment. Very strong monopsonists might be completely insulated from competition by other firms, and practically all investments in their labor force would be specific. On the other hand, firms in ex-

[13] To judge from a sample of firms recently analyzed, formal orientation courses are quite common, at least in large firms (see H. F. Clark and H. S. Sloan, *Classrooms in the Factories,* New York, 1958, Chap. IV).

tremely competitive labor markets would face a constant threat of raiding and would have fewer specific investments available.

These examples convey some of the surprisingly large variety of situations that come under the rubric of specific investment. This set is now treated abstractly in order to develop a general formal analysis. Empirical situations are brought in again after several major implications of the formal analysis have been developed.

If all training were completely specific, the wage that an employee could get elsewhere would be independent of the amount of training he had received. One might plausibly argue, then, that the wage paid by firms would also be independent of training. If so, firms would have to pay training costs, for no rational employee would pay for training that did not benefit him. Firms would collect the return from such training in the form of larger profits resulting from higher productivity, and training would be provided whenever the return—discounted at an appropriate rate—was at least as large as the cost. Long-run competitive equilibrium requires that the present value of the return exactly equals costs.

These propositions can be stated more formally with the equations developed earlier. According to equations (5) and (7), the equilibrium of a firm providing training in competitive markets can be written as

$$MP_0' + G\left[\sum_{t=1}^{n-1} \frac{MP_t - W_t}{(1+i)^t}\right] = W_0 + C, \qquad (11)$$

where C is the cost of training given only in the initial period, MP_0' is the opportunity marginal product of trainees, W_0 is the wage paid to trainees, and W_t and MP_t are the wage and marginal product in period t. If the analysis of completely specific training given in the preceding paragraph is correct, W would always equal the wage that could be received elsewhere, $MP_t - W_t$ would be the full return in t from training given in 0, and G would be the present value of these returns. Since MP_0' measures the marginal product elsewhere and W_0 measures the wage elsewhere of trainees, MP_0' equals W_0. As a consequence G equals C, or, in full equilibrium, the return from training equals costs.

Before claiming that the usual equality between marginal product and wages holds when completely specific training is considered, the reader should bear in mind two points. The first is that the equality between wages and marginal product in the initial period involves

opportunity, not actual marginal product. Wages would be greater than actual marginal product if some productivity was foregone as part of the training program. The second is that, even if wages equaled marginal product initially, they would be less in the future because the differences between future marginal products and wages constitute the return to training and are collected by the firm.

All of this follows from the assumption that firms pay all costs and collect all returns. But could not one equally well argue that workers pay all specific training costs by receiving appropriately lower wages initially and collect all returns by receiving wages equal to marginal product later? In terms of equation (11), W_t would equal MP_t, G would equal zero, and W_0 would equal $MP_0' - C$, just as with general training. Is it more plausible that firms rather than workers pay for and collect and return from training?

An answer can be found by reasoning along the following lines. If a firm had paid for the specific training of a worker who quit to take another job, its capital expenditure would be partly wasted, for no further return could be collected. Likewise, a worker fired after he had paid for specific training would be unable to collect any further return and would also suffer a capital loss. The willingness of workers or firms to pay for specific training should, therefore, closely depend on the likelihood of labor turnover.

To bring in turnover at this point may seem like a *deus ex machina* since it is almost always ignored in traditional theory. In the usual analysis of competitive firms, wages equal marginal product, and since wages and marginal product are assumed to be the same in many firms, no one suffers from turnover. It would not matter whether a firm's labor force always contained the same persons or a rapidly changing group. Any person leaving one firm could do equally well in other firms, and his employer could replace him without any change in profits. In other words, turnover is ignored in traditional theory because it plays no important role within the framework of the theory.

Turnover becomes important when costs are imposed on workers or firms, which are precisely the effects of specific training. Suppose a firm paid all the specific training costs of a worker who quit after completing it. According to our earlier analysis, he would have been receiving the market wage and a new employee could be hired at the same wage. If the new employee were not given training, his marginal product would be less than that of the one who quit since presumably

training raised the latter's productivity. Training could raise the new employee's productivity but would require additional expenditures by the firm. In other words, a firm is hurt by the departure of a trained employee because an equally profitable new employee could not be obtained. In the same way an employee who pays for specific training would suffer a loss from being laid off because he could not find an equally good job elsewhere. To bring turnover into the analysis of specific training is not, therefore, a *deus ex machina* but is made necessary by the important link between them.

Firms paying for specific training might take account of turnover merely by obtaining a sufficiently large return from those remaining to counterbalance the loss from those leaving. (The return on "successes"—those remaining—would, of course, overestimate the average return on all training expenditures.) Firms could do even better, however, by recognizing that the likelihood of a quit is not fixed but depends on wages. Instead of merely recouping on successes what is lost on failures, they might reduce the likelihood of failure itself by offering higher wages after training than could be received elsewhere. In effect, they would offer employees some of the return from training. Matters would be improved in some respects but worsened in others, for the higher wage would make the supply of trainees greater than the demand, and rationing would be required. The final step would be to shift some training costs as well as returns to employees, thereby bringing supply more in line with demand. When the final step is completed, firms no longer pay all training costs nor do they collect all the return but they share both with employees.[14] The shares of each depend on the relations between quit rates and wages, layoff rates and profits, and on other factors not discussed here, such as the cost of funds, attitudes toward risk, and desires for liquidity.[15]

[14] A. Marshall (*Principles of Economics*, 8th ed., New York, 1949, p. 626) was clearly aware of specific talents and their effect on wages and productivity: "Thus the head clerk in a business has an acquaintance with men and things, the use of which he could in some cases sell at a high price to rival firms. But in other cases it is of a kind to be of no value save to the business in which he already is; and *then his departure would perhaps injure it by several times the value of his salary*, while probably he could not get half that *salary elsewhere*." (My italics.) However, he overstressed the element of indeterminacy in these wages ("their earnings are determined . . . by a bargain between them and their employers, the terms of which are theoretically arbitrary") because he ignored the effect of wages on turnover (*ibid.*, fn. 2).

[15] The rate used to discount costs and returns is the sum of a (positive) rate measuring the cost of funds, a (positive or negative) risk premium, and a liquidity premium that is presumably positive since capital invested in specific training is very illiquid (see the discussion in section 2 of Chapter III).

If training were not completely specific, productivity would increase in other firms as well, and the wage that could be received elsewhere would also increase. Such training can be looked upon as the sum of two components, one completely general, the other completely specific; the former would be relatively larger, the greater the effect on wages in other firms relative to the firms providing the training. Since firms do not pay any of the completely general costs and only part of the completely specific costs, the fraction of costs paid by firms would be inversely related to the importance of the general component, or positively related to the specificity of the training.

Our conclusions can be stated formally in terms of the equations developed earlier. If G is the present value of the return from training collected by firms, the fundamental equation is

$$MP' + G = W + C. \tag{12}$$

If G' measures the return collected by employees, the total return, G'', would be the sum of G and G'. In full equilibrium the total return would equal total costs, or $G'' = C$. Let a represent the fraction of the total return collected by firms. Since $G = aG''$ and $G'' = C$, equation (12) can be written as

$$MP' + aC = W + C, \tag{13}$$

or

$$W = MP' - (1 - a)C.[16] \tag{14}$$

Employees pay the same fraction of costs, $1 - a$, as they collect in returns, which generalizes the results obtained earlier. For if training were completely general, $a = 0$, and equation (14) reduces to equation (10); if firms collected all the return from training, $a = 1$, and (14) reduces to $MP_0' = W_0$; and if $0 < a < 1$, none of the earlier equations is satisfactory.

A few major implications of this analysis of specific training are now developed.

[16] If G'' did not equal C, these equations would be slightly more complicated. Suppose, for example, $G'' = G + G' = C + n$, $n \geq 0$, so that the present value of the total return would be greater than total costs. Then $G = aG'' = aC + an$, and

$$MP' + aC + an = W + C,$$

or

$$W = MP' - [(1 - a)C - an].$$

Rational firms pay generally trained employees the same wage and specifically trained employees a higher wage than they could get elsewhere. A reader might easily believe the contrary—namely, that general training would command a higher wage relative to alternatives than specific training does, since, after all, competition for persons with the latter is apt to be weaker than for those with the former. This view, however, overlooks the fact that general training raises the wages that could be received elsewhere while (completely) specific training does not, so a comparison with alternative wages gives a misleading impression of the *absolute* effect on wages of different types of training. Moreover, firms are not too concerned about the turnover of employees with general training and have no incentive to offer them a premium above wages elsewhere because the cost of such training is borne entirely by employees. Firms are concerned about the turnover of employees with specific training, and a premium is offered to reduce their turnover because firms pay part of their training costs.

The part of specific training paid by employees has effects similar to those discussed earlier for general training: it is also paid by a reduction in wages during the training period, tends to make age-earnings profiles steeper and more concave, etc. The part paid by firms has none of these implications, since current or future wages would not be affected.

Specific, unlike general, training produces certain "external" effects, for quits prevent firms from capturing the full return on costs paid by them, and layoffs do the same to employees. These, however, are external *diseconomies* imposed on the employees or employers of firms providing the training, not external economies accruing to other firms.

Employees with specific training have less incentive to quit, and firms have less incentive to fire them, than employees with no training or general training, which implies that quit and layoff rates are inversely related to the amount of specific training. Turnover should be least for employees with extremely specific training and most for those receiving such general training that productivity is raised less in the firms providing the training than elsewhere (as, say, in schools). These propositions are as applicable to the large number of irregular quits and layoffs that continually occur as to the more regular cyclical and secular movements in turnover; in this section, however, only the more regular movements are discussed.

Consider a firm that experiences an unexpected decline in demand for its output, the rest of the economy being unaffected. The marginal

product of employees without specific training—such as untrained or generally trained employees—presumably equaled wages initially, and their employment would now be reduced to prevent their marginal productivity from falling below wages. The marginal product of specifically trained employees initially would have been greater than wages. A decline in demand would reduce these marginal products too, but as long as they were reduced by less than the initial difference with wages, firms would have no incentive to lay off such employees. For sunk costs are sunk, and there is no incentive to lay off employees whose marginal product is greater than wages, no matter how unwise it was, in retrospect, to invest in their training. Thus workers with specific training seem less likely to be laid off as a consequence of a decline in demand than untrained or even generally trained workers.[17]

If the decline in demand were sufficiently great so that even the marginal product of specifically trained workers was pushed below wages, would the firm just proceed to lay them off until the marginal product was brought into equality with wages? To show the danger here, assume that all the cost of and return from specific training was paid and collected by the firm. Any worker laid off would try to find a new job, since nothing would bind him to the old one.[18] The firm might be hurt if he did find a new job, for the firm's investment in his training might be lost forever. If specifically trained workers were not laid off, the firm would lose now because marginal product would be less than wages but would gain in the future if the decline in demand proved temporary. There is an incentive, therefore, not to lay off workers with specific training when their marginal product is only temporarily below wages, and the larger a firm's investment the greater the incentive not to lay them off.

A worker collecting some of the return from specific training would have less incentive to find a new job when temporarily laid off than others would: he does not want to lose his investment. His behavior while laid off in turn affects his future chances of being laid off, for

[17] A very similar argument is developed by Walter Oi in "Labor as a Quasi-fixed Factor of Production," unpublished Ph.D. dissertation, University of Chicago, 1961. Also, see his article with almost the same title in *Journal of Political Economy*, December 1962.

[18] Actually one need only assume that the quit rate of laid-off workers tends to be significantly greater than that of employed workers, if only because the opportunity cost of searching for another job is less for laid-off workers.

if it were known that he would not readily take another job, the firm could lay him off without much fear of losing its investment.

These conclusions can be briefly summarized. If one firm alone experienced an unexpected decline in demand, relatively few workers with specific training would be laid off, if only because their marginal product was initially greater than their wage. If the decline were permanent, all workers would be laid off when their marginal product became less than their wage and all those laid off would have to find jobs elsewhere. If the decline were temporary, specifically trained workers might not be laid off even though their marginal product was less than their wage because the firm would suffer if they took other jobs. The likelihood of their taking other jobs would be inversely related, and therefore the likelihood of their being laid off would be positively related, to the extent of their own investment in training.

The analysis can easily be extended to cover general declines in demand; suppose, for example, a general cyclical decline occurred. Assume that wages were sticky and remained at the initial level. If the decline in business activity were not sufficient to reduce the marginal product below the wage, workers with specific training would not be laid off even though others would be, just as before. If the decline reduced marginal product below wages, only one modification in the previous analysis is required. A firm would have a greater incentive to lay off specifically trained workers than when it alone experienced a decline because laid-off workers would be less likely to find other jobs when unemployment was widespread. In other respects, the implications of a general decline with wage rigidity are the same as those of a decline in one firm alone.

The discussion has concentrated on layoff rates, but the same kind of reasoning shows that a rise in wages elsewhere would cause fewer quits among specifically trained workers than among others. For specifically trained workers initially receive higher wages than are available elsewhere and the wage rise elsewhere would have to be greater than the initial difference before they would consider quitting. Thus both the quit and layoff rate of specifically trained workers would be relatively low and fluctuate relatively less during business cycles. These are important implications that can be tested with the data available.

Although quits and layoffs are influenced by considerations other than investment costs, some of these, such as pension plans, are more strongly related to investments than may appear at first blush. A

pension plan with incomplete vesting privileges[19] penalizes employees who quit before retirement and thus provides an incentive—often an extremely powerful one—not to quit. At the same time pension plans "insure" firms against quits for they are given a lump sum—the nonvested portion of payments—whenever a worker quits. Insurance is needed for specifically trained employees because their turnover would impose capital losses on firms. Firms can discourage such quits by sharing training costs and the return with employees, but they would have less need to discourage them and would be more willing to pay for training costs if insurance were provided. The effects on the incentive to invest in one's employees may have been a major stimulus to the development of pension plans with incomplete vesting.[20]

An effective long-term contract would insure firms against quits, just as pensions do and also insure employees against layoffs. Firms would be more willing to pay for all kinds of training—assuming future wages were set at an appropriate level—since a contract, in effect, converts all training into completely specific training. A casual reading of history suggests that long-term contracts have, indeed, been primarily a means of inducing firms to undertake large investments in employees. These contracts are seldom used today in the United States,[21] and while they have declined in importance over time, they were probably always the exception here largely because courts have considered them a form of involuntary servitude. Moreover, any enforcible contract could at best specify the hours required on a job, not the quality of performance. Since performance can vary widely, unhappy workers could usually "sabotage" operations to induce employers to release them from contracts.

Some training may be useful not in most firms nor in a single firm, but in a set of firms defined by product, type of work, or geographical location. For example, carpentry training would raise productivity primarily in the construction industry, and French legal training would not be very useful in the United States. Such training would tend to be paid by trainees, since a single firm could not readily col-

[19] According to the as yet unpublished National Bureau study of pensions, most plans have incomplete vesting.

[20] This economic function of incomplete vesting should caution one against conceding to the agitation for more liberal vesting privileges. Of course, in recent years pensions have also been an important tax-saving device, which certainly has been a crucial factor in their mushrooming growth.

[21] The military and the entertainment industry are the major exceptions.

lect the return,[22] and in this respect would be the same as general training. In one respect, however, it is similar to specific training. Workers with training "specific" to an industry, occupation, or country are less likely to leave that industry, occupation, or country than other workers, so their industrial, occupational, or country "turnover" would be less than average. The same result is obtained for specific training, except that a firm rather than an industry, occupation, or country is used as the unit of observation in measuring turnover. An analysis of specific training, therefore, is helpful also in understanding the effects of certain types of "general" training.

Although a discrepancy between marginal product and wages is frequently taken as evidence of imperfections in the competitive system, it would occur even in a perfectly competitive system where there is investment in specific training. The investment approach provides a very different interpretation of some common phenomena, as can be seen from the following examples.

A positive difference between marginal product and wages is usually said to be evidence of monopsony power; just as the ratio of product price to marginal cost has been suggested as a measure of monopoly power, so has the ratio of marginal product to wages been suggested as a measure of monopsony power. But specific training would also make this ratio greater than one. Does the difference between the marginal product and the earnings of major-league baseball players, for example, measure monopsony power or the return on a team's investment? Since teams do spend a great deal on developing players, some and perhaps most of the difference must be considered a return on investment (even if there were no uncertainty about the abilities of different players).[23]

Earnings might differ greatly among firms, industries, and countries and yet there might be relatively little worker mobility. The usual explanation would be that workers were either irrational or faced with formidable obstacles in moving. However, if specific[24] training

22 Sometimes firms cooperate in paying training costs, especially when training apprentices (see *A Look at Industrial Training in Mercer County, N. J.*, Washington, 1959, p. 3).

23 S. Rottenberg ("The Baseball Players' Labor Market," *Journal of Political Economy*, June 1956, p. 254) argues that the strong restrictions on entry of teams into the major leagues is prima-facie evidence that monopsony power is important, but the entry or threat of new *leagues*, such as have occurred in professional basketball and football, is a real possibility. And, of course, new teams have entered in recent years.

24 Specific, that is, to the firms, industries, or countries in question.

were important, differences in earnings would be a misleading estimate of what "migrants" could receive, and it might be perfectly rational not to move. For example, although French lawyers earn less than American lawyers, the average French lawyer could not earn the average American legal income simply by migrating to the United States, for he would have to invest in learning English and American law and procedures.[25]

In extreme types of monopsony, exemplified by an isolated company town, job alternatives for both trained and untrained workers are nil, and all training, no matter what its nature, would be specific to the firm. Monopsony combined with control of a product or an occupation (due, say, to antipirating agreements) converts training specific to that product or occupation into firm-specific training. These kinds of monopsony increase the importance of specific training and thus the incentive to invest in employees.[26] The effect on training of less extreme monopsony positions is more difficult assess. Consider the monopsonist who pays his workers the best wage available elsewhere. I see no reason why training should have a systematically different effect on the foregone earnings of his employees than of those in competitive firms and, therefore, no reason why specific training should be more (or less) important to him. But monopsony power as a whole, including the more extreme manifestations, would appear to increase the importance of specific training and the incentive for firms to invest in human capital.

2. *Schooling*

A school can be defined as an institution specializing in the production of training, as distinct from a firm that offers training in conjunction with the production of goods. Some schools, like those for barbers, specialize in one skill, while others, like universities, offer a large and diverse set. Schools and firms are often substitute sources of particular skills. This substitution is evidenced by the shift over time,

25 Of course, persons who have not yet invested in themselves would have an incentive to migrate, and this partly explains why young persons migrate more than older ones. For a further explanation, see the discussion in Chapter III; also see the paper by L. Sjaastad, "The Costs and Returns of Human Migration," *Investment in Human Beings*, pp. 80-93.

26 A relatively large difference between marginal product and wages in monopsonies might measure, therefore, the combined effect of economic power and a relatively large investment in employees.

for instance, in law from apprenticeships in law firms to law schools and in engineering from on-the-job experience to engineering schools.[27]

Some types of knowledge can be mastered better if simultaneously related to a practical problem; others require prolonged specialization. That is, there are complementary elements between learning and work and between learning and time. Most training in the construction industry is apparently still best given on the job, while the training of physicists requires a long period of specialized effort. The development of certain skills requires both specialization and experience and can be had partly from firms and partly from schools. Physicians receive apprenticeship training as interns and residents after several years of concentrated instruction in medical schools. Or, to take an example closer to home, a research economist spends not only many years in school but also a rather extensive apprenticeship in mastering the "art" of empirical and theoretical research. The complementary elements between firms and schools depend in part on the amount of formalized knowledge available: price theory can be formally presented in a course, while a formal statement of the principles used in gathering and handling empirical materials is lacking. Training in a new industrial skill is usually first given on the job, since firms tend to be the first to be aware of its value, but as demand develops, some of the training shifts to schools.

A student does not work for pay while in school but may do so after or before school, or during vacations. His earnings are usually less than if he were not in school since he cannot work as much or as regularly. The difference between what could have been and what is earned (including any value placed on foregone leisure) is an important indirect cost of schooling. Tuition, fees, books, supplies, and unusual transportation and lodging expenses are other, more direct, costs. *Net* earnings can be defined as the difference between actual earnings and direct school costs. In symbols,

$$W = MP - k, \tag{15}$$

where MP is actual marginal product (assumed equal to earnings) and k is direct costs. If MP_0 is the marginal product that could have been received, equation (15) can be written as

$$W = MP_0 - (MP_0 - MP + k) = MP_0 - C, \tag{16}$$

[27] State occupational licensing requirements often permit on-the-job training to be substituted for school training (see S. Rottenberg, "The Economics of Occupational Licensing," *Aspects of Labor Economics*, pp. 3-20).

where C is the sum of direct and indirect costs and where net earnings are the difference between potential earnings and total costs. These relations should be familiar since they are the same as those derived for general on-the-job training, which suggests that a sharp distinction between schools and firms is not always necessary: for some purposes schools can be treated as a special kind of firm and students as a special kind of trainee. Perhaps this is most apparent when a student works in an enterprise controlled by his school, which frequently occurs at many colleges.

Our definition of student net earnings may seem strange since tuition and other direct costs are not usually subtracted from "gross" earnings. Note, however, that indirect school costs are implicitly subtracted, for otherwise earnings would have to be defined as the sum of observed and foregone earnings, and foregone earnings are a major cost of high-school, college, and adult schooling. Moreover, earnings of on-the-job trainees would be net of *all* their costs, including direct "tuition" costs. Consistent accounting, which is particularly important when comparing earnings of persons trained in school and on the job, would require that earnings of students be defined in the same way.[28]

Regardless of whether all costs or merely indirect costs are subtracted from potential earnings, schooling would have the same kind of implications as general on-the-job training. Thus schooling would steepen the age-earnings profile, mix together the income and capital accounts, introduce a negative relation between the permanent and current earnings of young persons, and (implicitly) provide for depreciation on its capital. This supports my earlier assertion that an analysis of on-the-job training leads to general results that apply to other kinds of investment in human capital as well.

3. Other Knowledge

On-the-job and school training are not the only activities that raise real income primarily by increasing the knowledge at a person's command. Information about the prices charged by different sellers would enable a person to buy from the cheapest, thereby raising his command over resources; information about the wages offered by different firms would enable him to work for the firm paying the highest. In both examples, information about the economic system and about

[28] Students often have negative net earnings and in this respect differ from most on-the-job trainees, although at one time many apprentices also had negative earnings.

consumption and production possibilities is increased, as distinct from knowledge of a particular skill. Information about the political or social system—the effect of different parties or social arrangements—could also significantly raise real incomes.[29]

Let us consider in more detail investment in information about employment opportunities. A better job might be found by spending money on employment agencies and situation-wanted ads, by using one's time to examine want ads, by talking to friends and visiting firms, or in Stigler's language by "search."[30] When the new job requires geographical movement, additional time and resources would be spent in moving.[31] These expenditures constitute an investment in information about job opportunities that would yield a return in the form of higher earnings than would otherwise have been received. If workers paid the costs and collected the return, an investment in search would have the same implications about age-earnings profiles, depreciation, etc., as general on-the-job training and schooling, although it must be noted that the direct costs of search, like the direct costs of schooling, are usually added to consumption rather than deducted from earnings. If firms paid the costs and collected the return, search would have the same implications as on-the-job specific training.

Whether workers or firms pay for search depends on the effect of a job change on alternatives: the larger the number of alternatives made available by a change, the large (not the smaller) is the fraction of costs that have to be paid by workers. Consider a few examples. Immigrants to the United States have usually found many firms that could use their talents, and these firms would have been reluctant to pay the high cost of transporting workers to the United States. In fact immigrants have almost always had to pay their own way. Even a system of contract labor, which was seen to be a means of protecting firms against turnover, was singularly unsuccessful in the United

[29] The role of political knowledge is systematically discussed in A. Downs, *An Economic Theory of Democracy*, New York, 1957, and more briefly in my "Competition and Democracy," *Journal of Law and Economics*, October 1958.

[30] See G. J. Stigler, "Information in the Labor Market," *Investment in Human Beings*, pp. 94-105.

[31] Studies of large geographical moves—those requiring both a change in employment and consumption—have tended to emphasize the job change more than the consumption change. Presumably money wages are considered to be more dispersed geographically than prices.

States and has been infrequently used.[32] Firms that are relatively insulated from competition in the labor market have an incentive to pay the costs of workers coming from elsewhere since they have little to worry about in the way of competing neighboring firms. In addition, firms would be willing partly to pay for search within a geographical area because some costs—such as an employment agency's fee—would be specific to the firm doing the hiring since they must be repeated at each job change.

4. Productive Wage Increases

One way to invest in human capital is to improve emotional and physical health. In Western countries today earnings are much more closely geared to knowledge than to strength, but in an earlier day, and elsewhere still today, strength had a significant influence on earnings. Moreover, emotional health increasingly is considered an important determinant of earnings in all parts of the world. Health, like knowledge, can be improved in many ways. A decline in the death rate at working ages may improve earning prospects by extending the period during which earnings are received; a better diet adds strength and stamina, and thus earning capacity; or an improvement in working conditions—higher wages, coffee breaks, and so on—may affect morale and productivity.

Firms can invest in the health of employees through medical examinations, lunches, or avoidance of activities with high accident and death rates. An investment in health that increased productivity to the same extent in many firms would be a general investment and would have the same effect as general training, while an investment in health that increased productivity more in the firms making it would be a specific investment and would have the same effect as specific training. Of course, most investments in health in the United States are made outside firms, in households, hospitals, and medical offices. A full analysis of the effect on earnings of such "outside" investment in health is beyond the scope of this study, but I would like to discuss a relation between on-the-job and "outside" human investments that has received much attention in recent years.

When on-the-job investments are paid by reducing earnings during

[32] For a careful discussion of the contract-labor system in the United States, see C. Erickson, *American Industry and the European Immigrant, 1860-1885*, Cambridge, Mass., 1957.

the investment period, less is available for investments outside the job in health, better diet, schooling, and other factors. If these "outside" investments were more productive, some on-the-job investments would not be undertaken even though they were very productive by "absolute" standards.

Before proceeding further, one point needs to be made. The amount invested outside the job would be related to current earnings only if the capital market was very imperfect, for otherwise any amount of 'outside" investment could be financed with borrowed funds. The analysis assumes, therefore, that the capital market is extremely imperfect, earnings and other income being a major source of funds.[33]

A firm would be willing to pay for investment in human capital made by employees outside the firm if it could benefit from the resulting increase in productivity. The only way to pay, however, would be to offer higher wages during the investment period than would have been offered since direct loans to employees are prohibited by assumption. When a firm gives a productive wage increase—that is, an increase that raises productivity—"outside" investments are, as it were, converted into on-the-job investments. Indeed, such a conversion is a natural way to circumvent imperfections in the capital market and the resultant dependence of the amount invested in human capital on the level of wages.

The discussion can be stated more formally. Let W represent wages in the absence of any investment, and let a productive wage increase costing an amount C be the only on-the-job investment. Total costs to the firm would be $\pi = W + C$, and since the investment cost is received by employees as higher wages, π would also measure total wages. The cost of on-the-job training is not received as higher wages, so this formally distinguishes a productive wage increase from other on-the-job investments. The term MP can represent the marginal product of employees when wages equal W, and G the gain to firms from the investment in higher wages. In full equilibrium,

$$MP + G = W + C = \pi. \tag{17}$$

Investment would not occur if the firm's gain was nil ($G = 0$), for then total wages (π) would equal the marginal product (MP) when there is no investment.

[33] Imperfections in the capital market with respect to investment in human capital are discussed in section 2 of Chapter III.

It has been shown that firms would benefit more from on-the-job investment the more specific the productivity effect, the greater their monopsony power, and the longer the labor contract; conversely, the benefit would be less the more general the productivity effect, the less their monopsony power, and the shorter the labor contract. For example, a wage increase spent on a better diet with an immediate impact on productivity might well be granted,[34] but not one spent on general education with a very delayed impact.[35]

The effect of a wage increase on productivity depends on the way it is spent, which in turn depends on tastes, knowledge, and opportunities. Firms might exert an influence on spending by exhorting employees to obtain good food, housing, and medical care, or even by requiring purchases of specified items in company stores. Indeed, the company store or truck system in nineteenth century Great Britain has been interpreted as partly designed to prevent an excessive consumption of liquor and other debilitating commodities.[36] The prevalance of employer paternalism in underdeveloped countries has frequently been accepted as evidence of a difference in temperament between East and West. An alternative interpretation suggested by our study is that an increase in consumption has a greater effect on productivity in underdeveloped countries, and that a productivity advance raises profits more there either because firms have more monopsony power or because the advance is less delayed. In other

[34] The more rapid the impact, the more likely it is that it comes within the (formal or de facto) contract period. Leibenstein apparently initially assumed a rapid impact when discussing wage increases in underdeveloped countries (see his "The Theory of Underemployment in Backward Economies," *Journal of Political Economy*, April 1957). In a later comment he argued that the impact might be delayed ("Underemployment in Backward Economies: Some Additional Notes," *Journal of Political Economy*, June 1958).

[35] Marshall (*Principles of Economics*, p. 566) discusses delays of a generation or more and notes that profit-maximizing firms in competitive industries have no incentive to grant such wage increases.

"Again, in paying his workpeople high wages and in caring for their happiness and culture, the liberal employer confers benefits which do not end with his own generation. For the children of his workpeople share in them, and grow up stronger in body and in character than otherwise they would have done. The price which he has paid for labour will have borne the expenses of production of an increased supply of high industrial faculties in the next generation: but these faculties will be the property of others, who will have the right to hire them out for the best price they will fetch: neither he nor even his heirs can reckon on reaping much material reward for this part of the good that he has done."

[36] See G. W. Hilton, "The British Truck System in the Nineteenth Century," *Journal of Political Economy*, April 1957, pp. 246-247.

words "paternalism" may simply be a way of investing in the health and welfare of employees in underdeveloped countries.

An investment in human capital would usually steepen age-earnings profiles, lowering reported earnings during the investment period and raising them later on. But an investment in an increase in earnings may have precisely the opposite effect, raising reported earnings more during the investment period than later and thus flattening age-earning profiles. The cause of this difference is simply that reported earnings during the investment period tend to be net of the cost of general investments and gross of the cost of an increase in productive earnings.[37]

The productivity of employees depends not only on their ability and the amount invested in them both on and off the job but also on their motivation, or the intensity of their work. Economists have long recognized that motivation in turn partly depends on earnings because of the effect of an increase in earnings on morale and aspirations. Equation (17), which was developed to show the effect of investments outside the firm financed by an increase in earnings, can also show the effect of an increase in the intensity of work "financed" by an increase in earnings. Thus W and MP would show initial earnings and productivity, C the increase in earnings, and G the gain to firms from the increase in productivity caused by the "morale" effect of the increase in earnings. The incentive to grant a morale-boosting increase in earnings, therefore, would depend on the same factors as does the incentive to grant an increase used for outside investments. Many recent discussions of wages in underdeveloped countries have stressed the latter,[38] while earlier discussions often stressed the former.[39]

[37] If E represents reported earnings during the investment period and MP the marginal product when there is no investment, $E = MP - C$ with a general investment, $E = MP$ with a specific investment paid by the firm, and $E = MP + C$ with an increase in productive earnings.

[38] See Leibenstein, *Journal of Political Economy*, April 1957, and H. Oshima, "Underdevelopment in Backward Economies: An Empirical Comment," *Journal of Political Economy*, June 1958.

[39] For example, Marshall stressed the effect of an increase in earnings on the character and habits of working people (*Principles of Economics*, pp. 529-532, 566-569).

CHAPTER III

Investment in Human Capital: Rates of Return

THE most important single determinant of the amount invested in human capital may well be its profitability or rate of return, but the effect on earnings of a change in the rate of return has been difficult to distinguish empirically from a change in the amount invested. For since investment in human capital usually extends over a long and variable period, the amount invested cannot be determined from a known "investment period." Moreover, the discussion of on-the-job training clearly indicated that the amount invested is often merged with gross earnings into a single net earnings concept (which is gross earnings minus the cost of or plus the return on investment).

1. Relation Between Earnings, Costs, and Rates of Return

In this section, some important relations between earnings, investment costs, and rates of return are derived. They permit one to distinguish, among other things, a change in the return from a change in the amount invested. The discussion proceeds in stages from simple to complicated situations. First, investment is restricted to a single period and returns to all remaining periods; then investment is distributed over a known group of periods called the investment period. Finally, it is shown how the rate of return, the amount invested, and the investment period can all be derived from information on net earnings alone.

The discussion is from the viewpoint of workers and is, therefore, restricted to general investments; since the analysis of specific investments and firms is very similar, its discussion is omitted.

Let Y be an activity providing a person entering at a particular