Gary S. Becker

Human Capital

A Theoretical and Empirical Analysis,
with Special Reference to Education

Third Edition

This book is published by arrangement with the National Bureau of Economic Research

The University of Chicago Press, Chicago 60637
The University of Chicago Press, Ltd., London

Printed in the United States of America
02 01 00 99 98 97 96 95 2345
ISBN: 0-226-04119-0 (cloth)
0-226-04120-4 (paper)

Library of Congress Cataloging in Publication Data
Becker, Gary Stanley, 1930-
Human Capital : a theoretical and empirical analysis, with special reference to education / Gary S. Becker.-3rd ed.
p. cm.
Includes index.
1. Education-Economic aspects-United States. 2. Manpower policy-United States. 3. Human capital-United States. 1. Title.
LC66.B4 1993
331 .1 l'423-dc2O 93-24690
CIP

@The Paper used in this publication meets the minimum requirements of the American National Standard for Information Sciences-Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

Human Capital

CHAPTER I

Introduction to the Second Edition[1]

In the preface to the first edition, written about a decade ago, I remarked that in the preceding few years "interest in the economics of education has mushroomed throughout the world." The mushrooming has continued unabated; a bibliography on the economics of education prepared in 1957 would have contained less than 50 entries, whereas one issued in 1964 listed almost 450 entries and its second edition in 1970 listed over 1300 entries.[2] Moreover, this bibliography excludes the economic literature on health, migration, and other nonschooling investments in human capital, which has expanded even faster.[3]

This sustained interest in human capital and the continuing attention shown to the first edition of this book has encouraged me to issue a second edition. Nothing in the first edition has been changed; even the errors remain, conspicuous as they are to me now.[4] I have, however, incorporated three additional papers written within the first few years after the publication of the first edition. One of these three additions has not previously been published and another has not been readily available.

Chapter II developed an analysis of postschool investment and used it to explain age-earnings profiles and to interpret data on earnings per hour. That chapter also introduced a distinction between specific and general training to explain the relation between job skills and labor turnover, and the "hoarding" of labor during cyclical swings in business. These concepts have spawned a large and important literature that has successfully explained many aspects of the labor market in the United States and elsewhere.[5]

---

[1] I am indebted for helpful suggestions to Robert Michael, Victor Fuchs, and William Landes.

[2] See M. Blaug, *Economics* of *Education,* 2nd ed., London, 1970.

[3] I do not attempt to summarize or survey this growing body of literature on investments in human capital. A number of surveys and collections of essays have been published recently and the interested reader is referred to these. See, for example, UNESCO, *Readings in the Economics* of *Education,* United Nations Educational, Scientific and Cultural Organization, Paris, France, 1968; M. Blaug, *Economics* of *Education,* Elmsford, N.Y., 1970; B. F. Kiker, *Investment in Human Capital,* Columbia, S.C., 1971. Within the National Bureau of Economic Research, there have been three recent surveys of certain aspects of this literature; see Jacob Mincer, "The Distribution of Labor Incomes: A Survey with Special Reference to the Human Capital Approach," *Journal* of *Economic Literature,* 8, 1, March 1970; Finis Welch, "The NBER Approach to Human Resources Problems," *NBER Annual Report,* September *1971;* and Theodore *W.* Schultz, "Human Capital: *Policy* Issues and Research Opportunities," in *Human Resources,* Fiftieth Anniversary Colloquium, *Vol. VI,* NBER, *1972.*

[4] Let me mention only two here. In the adjustment (in Appendix A, section IC) to determine what earnings would have been if nobody had been unemployed, *I* used the duration of unemployment; this was incorrect because *I* had, and used, information on the fraction unemployed. *(I* am indebted to Robert *Solow* for pointing out this error.) Fortunately, a correct adjustment gives only slightly different results from the incorrect one used. There *is* a more serious error in illy discussion of the riskiness of investments in education (Chapter *IV,* section *4). 1* ignored the then developing literature on optimal portfolios, and did not derive my measure of marginal risk-the variance in the rate of return-from an analysis of utility maximization. *(I* am indebted to Lawrence Olson for pointing out these difficulties to me.)

[5] For a sampling of this literature, see Donald *0.* Parsons, "Specific Human Capital: An Application to Quit Rates and Layoff Rates," *Journal of Political Economy, 80, 6, 1120-1143* (November- December *1972);* Sherwin Rosen, "Learning and Experience in the Labor Market," *journal of Human Resources, 7, 3,* Summer *1972, pp. 326-342;* Lester Telser, *Competition, Collusion, and Game Theory,* Chicago, *1972;* Masatoshi

Chapter III introduced an analysis of the accumulation of human capital over the life cycle to explain, among other things, the shape of age-earnings profiles, the concentration of investments at earlier ages, and the personal distribution of earnings. This chapter also helped stimulate a large and empirically relevant literature.[6]

The personal distribution of earnings is partly determined by the distribution of, and the returns from, human capital. Mincer is responsible for the pioneering analysis that relates the distribution of earnings to human capital.[7] Section 3 of Chapter III extended his analysis by relating the distribution of earnings explicitly to rates of return and investment costs.

The additional material added in the second edition includes a portion of a paper, written jointly with Barry R. Chiswick,[8] which provides a convenient formulation for statistical estimation of the relation between the log of earnings, rates of return to human capital, and the time spent investing in human capital. Regression equations derived from this formulation are developed to estimate the contribution of schooling to earnings inequality. in the United States, especially its contribution to the difference in earnings inequality be tween the South and the North. This line of empirical analysis has more recently been extended to include postschool investment in a major study by Mincer,[9] and in other studies as well.[10]

In the first edition, although Chapter III assumed that individuals maximize their well-being as they accumulate human capital over their lifetime, no explicit model of utility or wealth maximization was developed. Therefore, the factors determining the distribution of investments at different ages were not explicitly analyzed. In my Woytinsky Lecture, published in 1967 and reprinted here as an addendum to Chapter III (see p. 94), a model of wealth maximization is developed that explains the distribution of investments, in particular the decline in investments over time, by (a) the decline in benefits from additional capital as fewer years of life remain, and (b) the rise in investment costs because foregone earnings rise as human capital is accumulated.[11]

Kuratani, "A Theory of Training, Earnings, and Employment: An Application to Japan," *Ph.D.* dissertation, Columbia University, *1973;* and L. Landes, "Male-Female Wage Differentials by Occupation," Ph.D. dissertation, Columbia University, *1973.*

[6] See, for example, Jacob Mincer, "On-the-job Training: Costs, Returns, and Some Implications," *Journal of Political Economy, 70, 5,* Part *2,* October *1962, pp. 50-79;* Yoram Bcn-Porath, "The Production of Human Capital and the Life Cycle of Earnings," *Journal of Political Economy, 75, 4,* August *1967,* Part *I, pp. 352*-365; Michael Grossman, "On the Concept of Health Capital and the Demand for Health," *Journal of Political Economy,* 80, 2, March-April 1972, pp. 223-255; and Yoram Weiss, "Investment in Graduate Education," *American Economic Review,* 61, December 1971, pp. 833-852.

[7] See Jacob Mincer, "Investment in Human Capital and Personal Income Distribution," *Journal of Political Economy,* August 1958.

[8] Gary S. Becker and Barry R. Chiswick, "Education and the Distribution of Earnings," *American Economic Review,* May 1966.

[9] See his *Schooling, Experience, and Earnings,* NBER, 1974.

[10] See, for example, Barry R. Chiswick, ***Income*** *Inequality: Regional Analyses within a Human Capital Framework,* NBER, 1974; Thomas Johnson, "Returns from Investment in Human Capital," *American Economic Review,* 60, 4, September 1970, pp. 546-560; C. Michael Rahm, "The Occupational Wage Structure," Ph.D. dissertation, Columbia University, 1971; Jacob Mincer and Solomon Polachek, "Family Investments in Human Capital: Earnings of Women," *Journal* of *Political Economy,* 82, 2, March-April 1974; and Frank Stafford and G. Johnson, "The Earnings and Promotion of Women Faculty," Department of Economics, University of Michigan, mimeo, February 1973.

[11] At about the same time, a similar but more rigorously formulated model was independently developed by Ben-Porath (op. cit.).

Here the analysis goes behind the distribution of human capital and rates of return and examines the underlying distribution of opportunities and abilities. Since the observed distribution of earnings results from the interaction of these underlying distributions, the relative importance of opportunities and abilities is not easily "identified," although some tests are suggested. I have added a supplement to this discussion of "identifiably" that is motivated by many recent attempts to assess the independent effect of family background on earnings. It shows why these attempts understate the effect of background, and overstate the effect of human capital, on earnings, perhaps by substantial amounts.

The Woytinsky lecture also analyzes the effects on inequality and skewness in earnings of more equal opportunity, minimum schooling legislation, and "objective" selection of applicants to scarce places in schools. In it I attempt to explain, too, why earnings are more equally distributed and less skewed than incomes from nonhuman capital. Although the formulation has some unsolved analytical difficulties, I believe that this paper opens up a promising line of investigation that has received insufficient attention.[12]

The models of capital accumulation in the lecture-and in BenPorath's paper and several subsequent ones-have several limitations. Since the total hours supplied to the market sector are taken as given, these models do not consider the interaction between changes in wage rates over the life cycle resulting from the accumulation of human capital and the optimal allocation of time between the market and nonmarket sectors. Moreover, human capital is assumed to affect only earnings and the production of additional human capital, and to have no direct effect on utility or consumption.

These and some other restrictions are relaxed in the final essay added to this second edition. This paper, which I wrote and circulated in 1967 but never published, builds on the new approach to household behavior. In this approach, households produce the commodities that enter their utility functions by combining market-purchased goods and services, their own time, and human capital and other environmental variables.[13] With this approach I consider the uses of an individual's time at different ages; in particular I focus on the allocation of time to three activities: the production of nonmarket commodities (nonmarket time); the production of human capital (investment time); and the production of earnings (labor market time). I am also able to treat systematically a direct effect of human capital on consumption by permitting it to affect the efficiency of household production.[14]

The empirical analysis from the first edition is left intact, even though a substantial body of additional evidence has been accumulated since then, because the major findings have stood up remarkably well to the additional evidence. These findings include:

1. The average money rate of return on a college education to white males is between 11 and 13 per cent, with higher rates on a high-school education, and still higher rates on an elementary-school education. This range for the rate of return on college education, as well

---

[12] However, see the discussions in Mincer, "The Distribution of Labor Incomes: A Survey with Special Reference to the Human Capital Approach," Journal of *Economic* Literature, 8, 1, March 1970, pp. 1-26; Barry Chiswick, "Minimum Schooling Legislation and the Cross-Sectional Distribution of Income," *Economic* journal, 79, 3.5, September 1969, pp. 495-507; and Sherwin Rosen, "Income Generating Functions and Capital Accumulation," Harvard Institute of Economic Research, June 1973, unpublished.

[13] The approach is developed in my "A Theory of the Allocation of Time," *Economic Journal,* September 1965. A recent exposition can be found in Robert T. Michael and Gary S. Becker, "The New Approach to Consumer Behavior," *Swedish Journal of Economics,* 75, 4, 1973.

[14] A more extensive treatment of this subject, including some empirical work, can be found in Robert T. Michael, *The Effect of Education on Efficiency in Consumption,* NBER, 1972.

as the decline in the rate with successive stages of schooling, has also been found in many subsequent studies.[15]

2. The higher earnings of, say, college graduates compared to highschool graduates are partly due to the college graduate's greater ability, ambition, health, and better educated and more successful parents. I concluded from an examination of several kinds of evidence that differences in these and related traits explain a relatively small part of the earnings differentials between college and high-school graduates (but a larger part of the differentials at lower education levels). Hence, rates of return to college graduates that are unadjusted for "selectivity" are not bad guides to the true rates. Subsequent studies have adjusted for selectivity with a variety of data sources, and their conclusions usually have been quite similar to mine.[16]

Several papers in recent years have tried to formalize the rather old notion that education is largely a device to screen out abler persons for employers, and that, therefore, only a small part of earnings differentials by education can be attributed to the education per se.[17] *Even if schooling also works in this way, the significance* of *private rates* of *return to education is not affected at all.* Moreover, it should be noted that virtually no effort has been made to determine the empirical importance of screening. Furthermore, several major empirical issues must be resolved if screening is to be the primary explanation of earnings differentials. For example, college would be a horrendously expensive "employment agency": each year of college cost a typical individual in 1970 at least $6000 and cost society at least $1500 more than that. Surely, a year on the job or a systematic and intensive interview and applicant-testing program must be a much cheaper and more effective way to screen. My own opinion is that schooling-as-screening must occur in a world with imperfect information, but is a relatively minor influence in determining earnings differentials by education.

3. The evidence I examined indicated that rates of return on college and high-school education declined from about 1900 to 1940, but not after 1940, even though the relative number of college and high-school graduates also grew rapidly after 1940. 1 concluded that demand shifted more toward educated persons after 1940, partly due to the rapid growth of expenditures on R. and D., military technology, and services. The absence of any decline in

---

[15] See, for example, W. L. Hansen, "Total and Private Rates of Return to Investment in Schooling," *Journal of Political Economy,* 71, April 1963, pp. 128-140; G. Hanoch, "An Economic Analysis of Earnings and Schooling," *Journal of Human Resources,* 2, Summer 1967, pp. 310-329; and T. W. Schultz, *Investment in Human Capital,* New York, 1971.

[16] For a sampling, see Orley Ashenfefter *et al.,* "Graduate Education, Ability, and Earnings," *Review of Economics and Statistics,* February *1968, pp. 78-86; Zvi* Griliches and *W.* M. Mason, "Education, Income, and Ability," *Journal of Political Economy, 80,* May-June *1972, pp. S74-SI03; W.* L. Hansen, B. A. Weisbrod, and *W.* J. Scanlon, "Schooling and Earnings of Low Achievers," *American Economic Review, 60, 3,* June *1970, pp. 409-418;* B. Weisbrod and *P.* Karpoff, "Monetary Returns to College Education, Student Ability and College Quality," *Review of Economics and Statistics,* November *1968;* and A. Leibowitz, "Home Investments in Children," *journal of Political Economy, 82, 2,* Supplement, March-April *1974, pp. S111-S131.*

[17] See *P.* J. Taubman and *T.* J. Wales, "Higher Education, Mental Ability, and Screening," *Journal of Political Economy, 8, 1,* January-February *1973, pp. 28-55;* M. Spence, "Market Signalling," Ph.D. dissertation, Department *of* Economics, Harvard University, *1972;* J. E. Stightz, "The Theory of 'Screening,' Education, and the Distribution of Income," Cowles Foundation Discussion Paper *#354,* Yale University, March *1973;* K. J. Arrow, "Higher Education as a Filter," in K. Lumsden, ed., *Efficiency in Universities,* New York, Elsevier, *1974.*

rates of return after 1940 has been confirmed in a few subsequent studies.[18] Perhaps the current (1973) weak market for highly skilled manpower is the beginning of a resumption of the earlier decline. Note, however, that the absence of any decline after 1940 is not unique in American history; skill differentials, and thus presumably rates of return on education, apparently did not decline from 1860 to 1890.[19]

4. Average money rates of return on education are not the same for all groups; they are higher on college education for urban white males than for black or rural males, and higher for black than for white women. The evidence I examined suggested that these differences in rates led to corresponding differences in the fraction of high-school graduates going on to college. This effect of rates of return on the incentive to acquire education has been found in other studies.[20] For example, a growth in the monetary return to blacks from a college education in the 1960s has apparently sizably increased their number going to college, as well as shifted their fields of specialization: out of professions that cater to segregated black markets, such as clergy and medicine, and into more integrated professions, such as business and engineering.[21]

5. In Chapter VII, I calculated age-human-wealth profiles for different education classes that show the relation between age and the present value of future earnings, and used them to understand, among other things, life-cycle variations in savings. Some studies have continued this analysis of the linkage between the accumulations of human and nonhuman wealth.[22] I also drew on evidence for slaves, the one example of an explicit market that trades and prices human capital stocks rather than simply the services yielded by these stocks. A major and insightful study has recently appeared that interprets the market for slaves in the United States in terms of the theory of investment in human capital.[23]

The continuing vigor of the research in human capital is increasing testimony that this area of study is not one of the many fads that pass through the economics profession, but an important and lasting contribution. The major reason, in my judgment, is that the theoretical and empirical analyses have been closely integrated, with the theory often inspired by empirical findings.[24] The intimate relation of theory and observation has built a strong foundation for future work that cannot easily be torn down or ignored.

Therefore, I am confident that the analysis of human capital will continue to be a fruitful field of research. Although important studies of the effects of human capital in the market sector can be expected, I anticipate that the excitement will be generated by studies of its effects in the nonmarket sector. Major insights into the determinants of fertility, the production of health, the benefits from schooling to women who do not participate in the labor force, the productivity of marriage, and other topics will result from an integration of

---

[18] See Z. Griliches, "Notes on the Role of Education in Production Functions and Growth Accounting," in *Education, Income and Human Capital, W.* L. Hansen, ed., NBER, 1970; and F. Welch, "Education in Production," *journal* of *Political Economy*, 78, 1, January- February 1970.

[19] See C. Long, *lVages and Earnings in the United States,* 1860-1890, Princeton, 1960.

[20] See R. B. Freeman, *The Market for College-Trained Manpower,* Cambridge, 1971.

[21] See R. B. Freeman, "Changes in the Labor Market for Black Americans, 1948-1972," Brookings Papers on Economic Activity 1, Washington, D.C., 1973, pp. 67-120; and Finis Welch, "Education and Racial Discrimination," in 0. Ashenfelter and A. Rees, eds., *Discrimination in Labor Markets,* Princeton, 1973.

[22] See G. Ghez and G. S. Becker, *The Allocation of Time and Goods over the Life Cycle,* NBER, 1974; and 1. Ehrlich and U. Ben-Zion, "A Model of Productive Saving," mimeo, University of Chicago, 1972.

[23] R. W. Fogel and S. Engerman, *Time on the Cross,* Boston, 1974

[24] By contrast, in some other areas of research, such as research on economic growth, much of the theory seems to have developed quite independently of any empirical studies.

the theory of human capital with the allocation of time, household production functions, and the theory of choice.[25]

In short, the prospects for the analysis of human capital look almost as bright to me today as they did during its salad days.

[25] For some beginnings, see Michael, op. *cit.;* Grossman, op. *cit.;* and the essays in T. W. Schultz, ed., *Economics of the Family: Marriage, Children, and Human Capital,* New York, NBER, 1975.

Introduction to the First Edition

Some activities primarily affect future well-being; the main impact of others is in the present. Some affect money income and others psychic income, that is, consumption. Sailing primarily affects consumption, on-the-job training primarily affects money income, and a college education could affect both. These effects may operate either through physical resources or through human resources. This study is concerned with activities that influence future monetary and psychic income by increasing the resources in people. These activities are called investments in human capital.

The many forms of such investments include schooling, on-the-job training, medical care, migration, and searching for information about prices and incomes. They differ in their effects on earnings and consumption, in the amounts typically invested, in the size of returns, and in the extent to which the connection between investment and return is perceived. But all these investments improve skills, knowledge, or health, and thereby raise money or psychic incomes.

Recent years have witnessed intensive concern with and research on investment in human capital, much of it contributed or stimulated by T. W. Schultz. The main motivating factor has probably been a realization that the growth of physical capital, at least as conventionally measured, explains a relatively small part of the growth of income in most countries. The search for better explanations has led to improved measures of physical capital and to an interest in less tangible entities, such as technological change and human capital. Also behind this concern is the strong dependence of modern military technology on education and skills, the rapid growth in expenditures on education and health, the age-old quest for an understanding of the personal distribution of income, the recent growth in unemployment in the United States, the Leontief scarce-factor paradox, and several other important economic problems.

The result has been the accumulation of a tremendous amount of circumstantial evidence testifying to the economic importance of human capital, especially of education. Probably the most impressive piece of evidence is that more highly educated and skilled persons almost always tend to earn more than others. This is true of developed countries as different as the United States and the Soviet Union, of underdeveloped countries as different as India and Cuba, and of the United States one hundred years ago as well as today. Moreover, few if any countries have achieved a sustained period of economic development without having invested substantial amounts in their labor force, and most studies that have attempted quantitative assessments of contributions to growth have assigned an important role to investment in human capital. Again, inequality in the distribution of earnings and income is generally positively related to inequality in education and other training. To take a final example, unemployment tends to be strongly related, usually inversely, to education.

Passions are easily aroused on this subject and even people who are generally in favor of education, medical care, and the like often dislike the phrase "human capital" and still more any emphasis on its economic effects. They are often the people who launch the most bitter

attacks on research on human capital, partly because they fear that emphasis on the "material" effects of human capital detracts from its 11 cultural" effects, which to them are more important. Those denying the economic importance of education and other investments in human capital have attacked the circumstantial evidence in its favor. They argue that the correlation between earnings and investment in human capital is due to a correlation between ability and investment in human capital, or to the singling out of the most favorable groups, such as white male college graduates, and to the consequent neglect of women, dropouts, nonwhites, or high-school graduates. They consider the true correlation to be very weak, and, therefore, a poor guide and of little help to people investing in human capital. The association between education and economic development or between inequality in education and income is attributed to the effect of income on education, considering education as a consumption good, and hence of no greater causal significance than the association between automobile ownership and economic development or between the inequality in ownership and incomes.

This study hopes to contribute to knowledge in this area by going far beyond circumstantial evidence and analysis. Part One treats the theory of investment in human capital in detail and reveals its importance through the wide variety of economic phenomena that it encompasses. Chapter II derives a number of important effects of such investments on earnings and employment, while Chapter III shows how to estimate the total amount invested and how it changes when the anticipated gains change.

Part Two presents various empirical tests of the theoretical analysis. Chapters IV and V estimate the gains from college education in the United States in recent years. Costs as well as returns are considered, and estimates are presented not only for selected groups, such as white male college graduates, but also for typical college entrants (sections I and 3 of Chapter IV). Detailed attention is paid to the effect of the correlation between education and ability, and to the variation in the gain from college (sections 2 and 4 of Chapter IV). Social as well as private gains are estimated, and both are compared to corresponding estimates for physical capital (Chapter V).

Chapter VI briefly extends the discussion to high-school education, considering social as well as private costs and returns, and the effect of differential ability (section 1). This chapter also tries to discover the secular trend in the United States during the twentieth century in the economic effects of high-school and college education (section 2).

Chapter VII tests the implications of the theoretical analysis concerning the effect of human capital on the shape of age-earnings profiles (section 1). Also considered is the effect on the relation between age and the discounted value of subsequent earnings, which are called age-wealth profiles. These profiles are applied to the study of lifecycle variations in savings and consumption, and in a few other ways (section 2).

Perhaps it is best to conclude the introduction by emphasizing that the attention paid to the economic effects of education and other human capital in this study is not in any way meant to imply that other effects are unimportant, or less important than the economic ones. The advantages of a division of labor are no less real here than they are in research in general. I would like to urge simply that the economic effects are important and have been relatively neglected, at least until recently.

CHAPTER II

Human Capital Revisited[1]

1. Introduction

A Ryerson lecturer is supposed to tell the audience what he or she has been doing to earn a living from the University. Therefore it is an appropriate occasion for me to review what is known about human capital, especially the progress during the quarter-century since I published a book with that title. What has been called the human capital "revolution" began about three decades ago. Its pioneers include Ted Schultz, Jacob Mincer, Milton Friedman, Sherwin Rosen, and several others associated with the University of Chicago.

To most of you, capital means a bank account, one hundred shares of IBM, assembly lines, or steel plants in the Chicago area (especially during a Ryerson lecture). These are all forms of capital in the sense that they yield income and other useful outputs over long periods of time.

But I am going to talk about a different kind of capital. Schooling, a computer training course, expenditures on medical care, and lectures on the virtues of punctuality and honesty are capital too in the sense that they improve health, raise earnings, or add to a person's appreciation of literature over much of his or her lifetime. Consequently, it is fully in keeping with the capital concept as traditionally defined to say that expenditures on education, training, medical care, etc., are investments in capital. However, these produce human, not physical or financial, capital because you cannot separate a person from his or her knowledge, skills, health, or values the way it is possible to move financial and physical assets while the owner stays put. This embodiment of human capital in people is depressingly illustrated by the reactions of Hong Kong residents to the takeover of Hong Kong in 1997 by China. Many local people are busy protecting against China's policies by selling off some of their local financial and physical assets in order to invest in safer foreign securities and property. At the same time, however, computer experts, top management, and other skilled personnel are leaving Hong Kong in droves to seek citizenship elsewhere. They cannot reduce the risk to their human capital from China by investing only part of the human capital abroad; they must go where their capital goes.

It may seem odd now, but I hesitated a while before deciding to call my book *Human Capital-and* even hedged the risk by using a long subtitle. In the early days, many people were criticizing this term and the underlying analysis because they believed it treated people like slaves or machines. My, how the world has changed! The name and analysis are now readily accepted by most people not only in all the social sciences, but even in the media. I was surprised when a few months ago *Business* Week magazine had a cover story titled "Human Capital." And more amazing still, this has been their most popular cover story in several decades.

However, I should add that the concept of human capital remains suspect within academic circles that organize their thinking about social problems around a belief in the exploitation of labor by capital. It is easy to appreciate the problems created for this view

[1] I appreciate the helpful comments of Guity Nashat, Sherwin Rosen, and George Stigler and the assistance of David Meltzer.

by the human capital concept. For if capital exploits labor, does human capital exploit labor too-in other words, do some workers exploit other workers? And are skilled workers and unskilled workers pitted against each other in the alleged class conflict between labor and capital? If governments are to expropriate all capital to end such conflict, should they also expropriate human capital, so that governments would take over ownership of workers as well?

You can see why an idea developed to understand the economic and social world has been thrust into ideological discussions. Yet the concept of human capital has been popular in Communist countries. My book and those by Schultz and others on human capital are extensively used in the Soviet Union, Eastern Europe, and China. Even before the recent reforms, economists and planners there had no trouble with the concept of investing capital in people.

I will try to avoid technical analysis and jargon, and concentrate on showing how the analysis of investments in human capital helps in understanding a large and varied class of behavior not only in the Western world, but also in developing countries and countries with very different cultures. My discussion follows modern economics and assumes that these investments usually are rational responses to a calculus of expected costs and benefits.

2. Education and Training

Education and training are the most important investments in human capital. My book showed, and so have many other studies since then, that high school and college education in the United States greatly raise a person's income, even after netting out direct and indirect costs of schooling, and after adjusting for the better family backgrounds and greater abilities of more educated people. Similar evidence is now available for many points in time from over one hundred countries with different cultures and economic systems. The earnings of more educated people are almost always well above average, although the gains are generally larger in less-developed countries. Consider the differences in average earnings between college and high school graduates in the United States during the past fifty years. After being reasonably stable at between 40 and 50 percent until the early 1960s, they rose during that decade and then fell rather sharply. This fall during the 1970s led some economists and the media to worry about "overeducated Americans" (see Freeman, 1976). The concept of human capital itself fell into some disrepute.

But as Kevin Murphy and Finis Welch document in a recent study (1989), the monetary gains from a college education rose sharply during the 1980s to the highest level during these fifty years. The earnings advantage of high school graduates over high school dropouts also increased. Talk about overeducated Americans has vanished, and it has been replaced by concern once more about whether the United States provides adequate quality and quantity of education and other training.

These concerns are stimulated by tough economic competition from a renewed Europe, Japan, Korea, and other Asian countries, by sluggish rates of productivity advance in the United States during the past fifteen years, by a large drop in SAT scores, and by the dismal performance of American high school students on international tests in mathematics.

For those who prefer a monetary bottom line, trends in the earnings of young persons in the United States provide good reason for concern about the preparation they are receiving. The trend has been disastrous for the 15 percent of all students and much larger percentage

of innercity blacks who fail to complete high school. Their real wage rates have fallen by more than 30 percent since the early 1970s. Whether because

of school problems, family instability, or other forces, young people without a college education are not being adequately prepared for work in modern economies.

A Labor Department commission on labor quality, of which I am a member, is considering what can be done to improve the quality of workers in the United States. The concerns that led to the creation of this commission have stimulated renewed academic interest in the analysis of human capital, which illustrates how research in social sciences responds, sometimes excessively, to public policy issues.

The fraction of high school graduates who entered college fell during the middle of the seventies when benefits from a college education dropped, and it rose again in the eighties when the benefits greatly in creased. This caused an unexpected boom in college enrollments during the past few years, despite the relatively few people who are reaching college age. So, alas, the large rise in applications to our College in recent years is not due solely to more widespread appreciation of the superb education it provides. Many educators expected enrollments in the eighties to decline not only or demographic reasons, but also because college tuition was rising rapidly. They were wrong because they failed to appreciate that benefits from college rose even faster than costs, and that high school graduates respond to changes in both benefits and costs.

One might believe that enrollments in college would be easy to predict since the number of persons graduating from high school can be predicted quite closely. But demographic-based college enrollment forecasts have been wide of the mark during the past twenty years, as Steve Stigler and 1, especially Steve, showed in a subcommittee report a few years ago to the Baker Commission. Such forecasts ignored the changing incentives to women, blacks, and older persons to enroll in college.

That human capital investments tend to respond rationally to benefits and costs is clearly indicated by changes in the education of women. Prior to the 1960s in the United States, women were more likely than men to graduate from high school but less likely to continue on to college. Women shunned math, sciences, economics, and law, and gravitated toward teaching, home economics, foreign languages, and literature. Since relatively few married women continued to work for pay, they rationally chose an education that helped in household production and no doubt also in the marriage market. All this has changed radically. The enormous increase in the participation of married women is the most important labor force change during the past twenty-five years. Many women now take little time off from their jobs even to have children. As a result, the value to women of market skills has increased enormously, and they are shunning traditional "women's fields" to enter accounting, law, medicine, engineering, and other subjects that pay well. Indeed, women now comprise one-third or so of enrollments in law, business, and medical schools, and many home economics departments have either shut down or are emphasizing the "new home economics," which is a true branch of economics.

The same trends in women's education are found in Great Britain, France, Scandinavia, Taiwan, Japan, Mexico, and other countries with large increases in the labor force participation of women, even when attitudes toward women differ greatly from those now prevalent in Europe and the United States. Whenever the labor force participation of married women has increased sharply, changes in the gains from work for pay have had a

more powerful effect on the behavior of women than have traditional -ideas about the proper role of women.

job opportunities for women at first improved slowly as they started to move up in business and the professions during the past several decades. But the trend accelerated sharply after the late 1970s. The ratio of the earnings of full-time working women and men has increased more rapidly since 1979 than during any previous period in our history, and women are becoming much more prominent in many highly skilled jobs. Improvements in the economic position of black women have been especially rapid, and they now earn just about as much as white women.

Although the civil rights movement clearly contributed to greater job opportunities for women and other minorities, it is far from the whole story. This can be seen from the fact that women progressed most rapidly under the Reagan administration, which was opposed to affirmative action and did not have an active Civil Rights Commission. In my judgment, women advanced primarily because of their greater attachment to the labor force. This in turn was stimulated by a large decline in fertility, a rapid increase in divorce, and the growing importance of the service sector. Human capital analysis assumes that schooling raises earnings and productivity mainly by providing knowledge, skills, and a way of analyzing problems. An alternative view, however, denies that schooling does much to improve productivity, and instead it stresses "credentialism"-that degrees and education convey information about the underlying abilities, persistence, and other valuable traits of people. According to extreme versions of this line of analysis, earnings of, for example, college graduates exceed those of high school graduates not because college education raises productivity, but because more productive students go on to college.

Credentialism obviously exists. But many kinds of evidence suggest that credentialism does not explain most of the positive association between earnings and schooling.

The main problem with credentialisin is that companies do not want information on success at schoolwork, but on abilities and performance in the context of working life: the discipline imposed by factories, the need to please customers and get along with fellow employees, and so forth. Success in the flexible, individualistic, and rather undisciplined university atmosphere in most countries and in high schools in the United States does not convey much relevant information. I tell my classes that eccentrics and nuts can last much longer as students than as workers, and they respond that the same is true of professors.

A cheaper and more efficient way to provide information to employers is for teenagers to enter directly into the labor force, as they did prior to the industrial revolution. Far more would be learned about their work-related abilities and other characteristics after six years of work experience than after six additional years of schooling. High school and college education has spread extensively in modern economies because the additional knowledge and information acquired in school is so important in technologically advanced economies. I should add that advocates of the credentialism approach have become rather silent in recent years with the growing concerns about schools and labor quality in the United States.

Of course, learning and training also occur outside of schools, especially on jobs. Even college graduates are not well prepared for the labor market when they leave school, and they are fitted into their jobs through formal and informal training programs. The amount of on-the job training ranges from an hour or so at simple jobs like dishwashing to several years at complicated tasks like engineering in an auto plant. The limited information available indicates that on-the-job training is an important source of the very large increase in earnings as workers gain greater experience at work. And recent bold estimates by Jacob

Mincer suggest that the total investment in on-the-job training may be almost as large as the investment in education.

After a few years of frequent job changes, most workers settle down and remain with the same company for a long time. Workers and their employers get bonded together in large part because of the on-the-job learning and training. Therefore, it is not surprising that job changes are common among unskilled workers and uncommon among skilled workers. It also appears that job changes are much less frequent in Japan than in the United States mainly because on-the-job investments in workers are greater in Japan. My friends in the humanities like Dick Stern may complain that so far I have only mentioned "money," or they might say "mere money. " Is there any place in human capital theory for education to appreciate literature, culture and the good life? Fortunately, nothing in the concept of human capital implies that monetary incentives need be more important than cultural and no monetary ones.

Obviously, it is much easier to quantify the monetary side, but, nevertheless, progress has been made on other aspects. Many studies show that education promotes health, reduces smoking, raises the propensity to vote, improves birth control knowledge, and stimulates the appreciation of classical music, literature, and even tennis. In an ingenious study that relies heavily on economic theory, Bob Michael (1972) quantifies some non-monetary benefits of education. His results and those of others indicate that such benefits of schooling are quite large, although for most people they are apparently smaller than monetary benefits.

3. Human Capital and the Family

No discussion of human capital can omit the influence of families on the knowledge, skills, values, and habits of their children. Parents who severely beat their children cause lasting damage, while at the other end of the spectrum, sympathetic and firm parents help motivate their children.

Large differences among young children grow over time with age and schooling because children learn more easily when they are better prepared. Therefore, even small differences among children in the preparation provided by their families are frequently multiplied over time into large differences when they are teenagers. This is why the labor market cannot do much for school dropouts who can hardly read and never developed good work habits, and why it is so difficult to devise policies to help these groups.

Parents have a large influence on the education marital stability, and many other dimensions of their children's lives. The term "underclass" describes families in which low education, welfare dependence, early pregnancy, and marital instability pass from parents to children. In light of this, it is rather surprising that although earnings of parents and children are positively related, the relation is not strong. For example, if parents' earnings in the United States are 20 percent above the mean of their generation, the children's earnings tend to be less than 6 percent above the mean of their own generation. Earnings of parents and children appear to be a little more strongly related when parents are poorer.

It is easy to see why children's and parents' earnings may be closer in poorer families. Richer families can pay for the training of their children, including the earnings foregone when children spend time in training rather than at work. Many poorer parents would be willing to lend their children money to help them obtain further training if the parents could expect to get paid back later when they are old. But children may not carry out their part of

the bargain, especially in highly mobile societies where children often live far from their parents.

One solution is for governments to lend money to students when their parents are unable or unwilling to finance the training. The federal government has developed an extensive loan program to help students finance college education. Unfortunately the program has serious flaws, including low caps on the maximum amounts that can be borrowed, misplaced and excessive subsidies, and shockingly high default rates. In addition to explicit loans, some direct subsidies to schools may, in effect, also be "loans" to students which they repay later with taxes that help finance support for the elderly. By combining publicly subsidized schooling with a social security system, countries may have found a very crude and indirect, but perhaps reasonably effective, way to provide loans to children that get repaid when the parents are old and collect retirement benefits (see Becker and Murphy, 1988).

Families divide their total spending on children between number of children and the amount spent per child. The number of children and spending per child tend to be negatively related. The reason is simple. An increased number of children raises the effective cost of adding to the spending on each child, because an additional dollar or hour of time spent on each child then means a larger total addition to spending. Similarly, an increase in the dollars or time spent on each child raises the cost of having an additional child. Consequently, even a modest tax on births can have a large negative effect on the number of children and a large positive effect on the amount spent on each child.

China imposed heavy, not modest, taxes and other penalties on large families during the past decade, especially in urban areas. It is revealing about the cross-cultural relevance of this analysis that sharp declines in urban fertility have been accompanied by discussions in the Chinese press of the "emperor child." This refers to only children who receive lavish toys and presents from their parents, and are pushed toward outstanding educational achievement.

This negative relation at the family level between number of children and spending per child implies a close and also usually negative relation at the aggregate level between population growth and investments in human capital. Differences among ethnic groups in the United States are fascinating. Groups with small families generally spend a lot on each child's education and training, while those with big families spend much less. The Japanese, Chinese, Jews, and Cubans have small families and the children become well educated, while Mexicans, Puerto Ricans, and blacks have big families and the education of children suffers. (I should add that the Mormons are an interesting exception, for they have both very large families and high levels of achievement). It should come as no surprise that children from the ethnic groups with small families and large investments in human capital typically rise faster and further in the United States' income-occupation hierarchy than do children from other groups.

Malthus' famous prediction that people marry earlier and birth rates rise when incomes increase was decisively contradicted by the industrial revolution, whose effects became evident only shortly after publication of the second edition of his book on population. This is a common paradox: a great book gets contradicted by events not long after publication. The contradiction to Malthus' theory is that fertility fell sharply, rather than rose, as per capita incomes grew in Great Britain, the United States, France, Germany, Sweden, and other Western countries. Rapid advances in education and other training accompanied the sharp declines in fertility. Parents did spend more on children when their incomes rose-as

Malthus predicted-but they spent a lot more on each child and had fewer children, as human capital theory predicts.

Similar changes occur in other cultures when they experience rapid economic growth. Taiwan's birth rate was cut in half from 1960 to 1975, while the fraction of high school graduates doubled after Taiwan took off in the 1960s toward its remarkable economic growth. Mexico's birth rate did not fall much during its rapid economic growth in the 1950s and 1960s. But since 1975 birth rates have fallen by more than one-third, and school enrollments have expanded rapidly.

4. Human Capital and Economic Development

Economic analysis has no trouble explaining why, throughout history, few countries have experienced very long periods of persistent growth in income per person. For if per capita income growth is caused by the growth of land and physical capital per worker, diminishing returns from additional capital and land eventually eliminate further growth. The puzzle, therefore, is not the lack of growth, but the fact that the United States, Japan, and many European countries have had continuing growth in per capita income during the past one hundred years and longer.

Presumably, the answer lies in the expansion of scientific and technical knowledge that raises the productivity of labor and other inputs in production. The systematic application of scientific knowledge to production of goods has greatly increased the value of education, technical schooling, and on-the-job training as the growth of knowledge has become embodied in people-in scientists, scholars, technicians, managers, and other contributors to output.

It is clear that all countries which have managed persistent growth in income have also had large increases in the education and training of their labor forces. First, elementary school education becomes universal, then high school education spreads rapidly, and finally children from middle income and poorer families begin going to college. A skeptic might respond that the expansion in education as countries get richer no more implies that education causes growth than does a larger number of dishwashers in richer countries imply that dishwashers are an engine of growth.

However, even economists know the difference between correlation and causation, and have developed rather straightforward methods for determining how much of income growth is caused by a growth in human capital. In an excellent study for the United States, Edward Denison (1985) finds that the increase in schooling of the average worker between 1929 and 1982 explains about one-fourth of the rise in per capita income during this period. He is unable to explain much of the remaining growth. I like to believe that this is mainly because he cannot measure the effects on earnings of improvements over time in health, on-the-job training, and other kinds of human capital.

The outstanding economic records of Japan, Taiwan, and other Asian economies in recent decades dramatically illustrate the importance of human capital to growth. Lacking natural resources- e.g., they import practically all their sources of energy-and facing discrimination from the West, these so-called Asian tigers grew rapidly by relying on a well trained, educated, hard-working, and conscientious labor force. It surely is no accident, for example, that Japan’s system of lifetime employment at large companies originated after World War II when they began to upgrade their technology rapidly partly by investing heavily in the training of employees. The lifetime system is not explained just by the tradi-

tional Japanese culture that emphasizes loyalty toward groups, for job changes in Japan were frequent during the first half of this century (see Hashimoto and Raisian, 1985).

Compelling evidence of the link between human capital and technology comes from agriculture. Education is of little use in traditional agriculture because farming methods and knowledge are then readily passed on from parents to children. Farmers in countries with traditional economies are among the least educated members of the labor force. By contrast, modern farmers must deal with hybrids, breeding methods, fertilizers, complicated equipment, and intricate futures markets for commodities. Education is of great value since it helps farmers adapt more quickly to new hybrids and other new technologies (see Welch, 1970). Therefore, it is no surprise that farmers are about as well educated as industrial workers in modern economies.

Education and training is also helpful in coping with changing technologies and advancing productivity in the manufacturing and service sectors. Recent studies show that more rapidly progressing industries do attract better-educated workers and provide greater training on the job (see Mincer and Higuchi, 1988; Gill, 1989).

5. Conclusions

We have reached the end of my visit. Perhaps I have succeeded in conveying the enormous energy devoted to the analysis of human capital during the past quarter-century and the impressive advances of analytical techniques and the accumulation of empirical regularities. Much is now known for many countries about the effects of education on earnings, occupation, employment, and unemployment of both men and women and various races and ethnic groups. Much too is known about the link between birth rates and investments in education and training, how families influence the human capital of their children, and the relation between investments in human capital and economic progress.

I indicated earlier that human capital analysis has been motivated partly by a desire to evaluate proposals to improve the quality of the work force through schooling, training, medical services, and child care. But its main purpose as far as I am concerned is to remove a little of the mystery from the economic and social world that we live in.

References

Becker, Gary S., and Kevin M. Murphy. "The Family and the State." *Journal of Law and Economics 31 (1988): 1-18.*

Denison, Edward E *Trends in American Economic Growth, 1929-1982.* Washington, D.C.: The Brookings institution, *1985.*

Freeman, Richard. *The Overeducated American.* New York: Academic Press, *1976.*

Gill, Indermit. "Technological Change, Education, and Obsolescence of Human Capital: Some Evidence for the U.S." University of Chicago, 1989.

Hashimoto, Masanori, and John Raisian. "Employment Tenure and Earnings Profiles in Japan and the United States." *American Economic Review* 75 (1985): 721-35.

Michael, Robert T. *The Effect of Education on Efficiency in Consumption.* New York: National Bureau of Economic Research, 1972.

Mincer, Jacob, and Yoshio Higuchi. "Wage Structures and Labor Turnover in the U.S. and Japan. " Journal *of the Japanese and International Economies* (1988): 297-331.

Murphy, Kevin M., and Finis Welch. "Wage Premiums for College Graduates: Recent Growth and Possible Explanations." *Educational Researcher* 18 (1989): 17-27.
Welch, Finis. "Education in Production. " Journal *of Political Economy* 78 (1970): 35-59.

Part One

Theoretical Analysis

"The most valuable of all capital is that invested in human beings."

Alfred Marshall, *Principles of Economics*

CHAPTER III

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 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

---

[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. 949.

[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, Harper, 1937; 1. 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.

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

$$MP = W, \quad (1)$$

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

$$MP_t = W_t , \quad (2)$$

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 receipts 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 *Et* and R, represent expenditures and receipts during period *t,* and i the market discount rate, then the equilibrium condition can be written as 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.

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

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 where *k* measures the outlay on training.

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

If a new term is defined,

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

equation (4) can be written as

$$MP_o + G = W_o + k. \quad (6)$$

Since the term *k* measures only 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

$$MP_o' + G = W_o + C. \quad (7)$$

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, must equal Wt for all $t = 1,\ldots n - 1$, and therefore

[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).

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

Equation (7) is reduced to

$$MP_o' = W_o + C. \qquad (9)$$

or

$$W_o = MP_o' - C. \qquad (10)$$

In terms of actual marginal product

$$MPo = Wo + k, \qquad (9')$$

or

$$Wo = MPo - k. \qquad (10')$$

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 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.

---

[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); 1 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," *Econometrical,* January 1955.

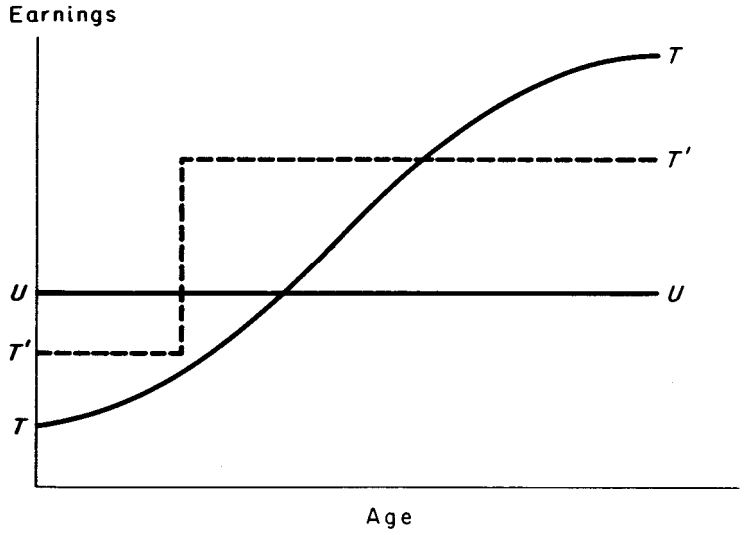
[8] R. Goode has argued (see his "Educational Expenditures and the Income Tax," in Selma J. Mushkin, ed., *Economics of Higher Education,* Washington, 1962) that educated persons should be permitted to subtract from income a depreciation allowance on tuition payments. Such an allowance is apparently 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.

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

CHART I

Relation of Earnings to Age

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

[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

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 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 reenlistment 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 reenlistments.

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 external economies to competitors. This argument and analogy would apply if firms were to pay training costs, for they would suffer a 11 capital loss"

---

other areas. A shortened version was published with the same title in the *Economic journal* of September 1965.

[11] See *Manpower Management and Compensation,* report of the Cordiner Committee, Washington, D.C., 1957, *VoL 1*, 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 an excellent study, see Gorman C. Smith, "Differential Pay for Military Technicians," unpublished Ph.D. dissertation, Columbia University, 1964.

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 may have difficulty establishing property rights in innovations, and these innovations may 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 unwonted 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 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

---

[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 and health, as well as on-the-job training. Protection would serve the same purpose as the creation of monopolies 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

[13] To judge from a sample of firms 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, Chapter IV).

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 extremely 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 equal 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

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

where C is the cost of training given only in the initial period, MP_0' is the opportunity marginal product of trainees, W_O is the wage paid to trainees, and *Wt* and *MPt* 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, *MPt - Wt* 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, 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 were 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, would equal *MP,* G would equal zero, and W_O would equal $MP_0' - C$, just as with general training. Is it More plausible that firms rather than workers pay for and collect any 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 introducing a *deus ex machina,* since turnover 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 the training. 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, to introduce 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 11 successes"-those remaining-would, of course, overestimate the average return on all training expenditures.) Firms could do even better, how-ever, 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 other, 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]

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

$$MP'+G= W+C. \qquad (12)$$

If G' measures the return collected by employees, the total return, G11, 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

$$MY+ aC = W + C, \qquad (13)$$

or

[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).

$$W = MP' - (1 - a)C.^{16} \qquad (14)$$

Employees pay the same fraction of costs, I - 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 $MP_O' = W_O$; 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.

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 (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

---

[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 c9sts. Then G = aG'' = aC + an, and

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

or

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

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 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

---

[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.

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 Jay off specifically trained workers than when it alone experience 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. 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 ~haring 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 enforceable 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

---

[19] According to the National Bureau study of pensions, most plans have incomplete vesting. See R. F. Murray, *Economic Aspects of Pensions: A Summary Report,* New York, NBER, 1968.

[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

trainees, since a single firm could not readily collect 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 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

---

[22] Sometimes firms cooperate in paying training costs, especially when training apprentices (see R. F. Arnold, A Look at Industrial Training in Mercer County, NJ., Washington, D.C., 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, are 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.

[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.*

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 to 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, 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,

---

[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.

[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).*

$$W = MP - k, \qquad (15)$$

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

$$W = MPo - (MPo - MP + k) = MPo - C, \qquad (16)$$

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 11 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 consumption and production possibilities is increased, as distinct from knowledge of a particular skill. Information about the political or social

[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.

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 larger (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 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

---

[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 111ore than the consumption change. Presumably money wages are considered to be more dispersed geographically than prices.

[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 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 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 11 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,

$$MP + G = W + C = \pi. \quad (17)$$

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

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.

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 prevalence 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 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.

---

[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.

[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.

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 discussions of wages in underdeveloped countries have stressed the latter,[38] while earlier discussions often stressed the former.[39]

[38] See Leibenstein, *Journal of Political Economy*, April *1957*, and *H.* Oshima, "Underdevelopment in Backward Economics: 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).*