SCHOOLING, EXPERIENCE, HOURS OF WORK,

AND EARNINGS IN CANADA

by

RICHARD DONALD SCOTT

B.A., Simon Fraser University, 1971

M.A., University of British Columbia, 1973

9 1

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

DEPARTMENT OF ECONOMICS

We accept this dissertation as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

September, 1979

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the Head of my Department or by his representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of ECONOMICS

The University of British Columbia
2075 Wesbrook Place
Vancouver, Canada
V6T 1W5

Date Sept 21, 1979.

ABSTRACT

This study investigates a broad range of factors which might be thought to influence the employment earnings of Canadian males. Micro-data drawn from the 1971 census are analysed, using as a frame of reference the human-capital model derived, and implemented for the United States, by Jacob Mincer.

Opening discussion furnishes a detailed critique of the model itself, and of the auxiliary hypotheses required to make it perform empirically. Particular emphasis is laid upon the implicit assumption of perpetual long-run equilibrium and upon the neglect of variables arising on the demand side of the labour market. Generally, it is argued that although the human-capital paradigm may serve as a framework for empirical description, it is inadequate as a scientific theory because it fails to generate a wide array of hypotheses which are clearly susceptible to falsification.

Earnings functions are estimated by ordinary least squares for a sample of almost 23,000 out-of-school males who worked, mainly in the private sector, at some time during 1970. Results yielded for Canada by the human-capital specification are compared with those reported by Mincer. The regressions are then expanded to include variables such as industry, region, and occupation, together with other personal attributes. These are found to rival the importance of the orthodox human-capital

variables. Contrary to United States results, the elasticity of earnings with respect to weeks (or hours) worked is less than unity.

In light of recent analyses which make human-capital investment and labour supply objects of simultaneous decision within a life-cycle context, further investigation is carried out using a simplified, two-equation, linear model in which earnings and hours are both endogenous. Estimates performed by the method of three-stage least squares indicate an elasticity of earnings with respect to hours considerably in excess of unity. However, within particular regional and industrial categories, wages and hours tend to be offsetting. Schooling coefficients, or "rates of return," fall in the 5.25-6.50% range.

Terence J. Wales

Research Supervisor

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENT

I should like to thank all the members of my committee, but especially its chairman, Terence Wales, who dispensed much patience and congeniality along with helpful substantive comment. I am equally indebted to Jonathan Kesselman, who supervised the present research at its early and intermediate stages. For excellent programming assistance I am grateful to the staff of the University of British Columbia Statistics Centre--in particular, Frank Flynn, Lewis James, and prior to his departure, Keith Wales. My task in preparing this final draft was considerably eased, through the competence and experience of the typist Maryse Ellis. Though the preceding individuals contributed a number of improvements, they bear no responsibility for any errors or omissions which may remain.

This study was carried out in part while I was in receipt of a Canada Council Doctoral Fellowship.

R. D. S.

INTRODUCTION

Owing to a scarcity of fertile data, Canadian research in the area of human capital has been limited, both in volume and scope.[1] As a consequence, we have had to glean, mainly from the American literature, most of what we presently know and teach, about the rates of return to investment in education, and about the complicated web of interaction linking such key variables as schooling, on-the-job training, hours of work and the level of individual earnings. The investigation reported here is an attempt to narrow the current research deficit. Results of this work supply a new description of the forces determining employment incomes in Canada, and at the same time, illuminate some important differences between Canadian and American experience.[2]

The present study selects as a point of departure the human-capital model of income determination, developed over the past two decades by a group of well-known economists, but consistently applied in its most uncompromising form by one member of the school, namely, Jacob Mincer.[3] With the publication of Mincer's recent book, *Schooling, Experience, and Earnings,*[4] human-capital orthodoxy appears to have reached a major empirical plateau. When fully deployed, Mincer's version of the human-capital model succeeds in accounting for just over half the variance of earnings in a large body of microdata drawn from the United States Census. In the process, it yields new estimates of the private return to investment in formal education and on-the-job training.

Until recently, empirical work of the kind reported by Mincer has been very difficult to pursue in Canada: except in a few special instances,[5] researchers have been without access to microdata. The decision by Statistics Canada to issue a large public file of individual observations drawn from the 1971 Census was therefore a welcome advance. Microdata extracted from this new and comparatively rich source, the so-called Public Use Sample, provides an empirical footing for the work reported here.[6]

The initial chapters of this dissertation concern the application of Mincer's theory and his empirical methods to the Canadian census data. Chapter I introduces the main theoretical arguments of the human-capital school and offers a critical appraisal. It is argued that the human-capital analysis fails to generate an adequate set of testable hypotheses, though it may serve as a convenient framework for empirical description. Chapter II considers various problems of implementation, since empirical measurements, even if only descriptive, may harbour misleading biases.

Chapter III exhibits two sets of regression equations. The first set replicates, as nearly as convenience and the data will allow, Mincer's human capital "earnings functions." On the one hand, this exercise furnishes some interesting comparative results for the Canadian economy, and on the other, serves the worthwhile scientific purpose of confronting the human-capital model with new data. The fact that Canadian and American results differ at some key points *without invalidating the model* supports the present contention that the standard theory is virtually immune from scientific falsification.

The second set of regressions in Chapter III explores the consequences of adding to the empirical model variables typically ignored by human-capital theorists. Among the variables inserted are dummies representing region, industry, occupation, urban residence, official language, ethnic and religious group, period of immigration, and family status. The resulting estimates, it is argued, provide a better basis for assessing the contribution of the "orthodox" variables than do Mincer's highly parsimonious specifications.

Although the task of replicating Mincer's work, and of exploring some alternative hypotheses with Canadian data, is in itself a substantial research undertaking, one seemingly important weakness in the application of the model invites a further stage of inquiry. The difficulty in question arises from Mincer's casual introduction of weeks worked as an exogenous variable in the earnings function. If weeks worked depend on the wage rate, and hence, upon earnings, by way of the individual's labour-supply response, including weeks worked on the right-hand side of a regression in which earnings are the dependent variable will necessarily bias the estimation. Moreover, the coefficients which Mincer and others interpret as rates of return will in fact be complex, displaying the tangled structural effects of both human-capital investment and labour supply (not to mention labour demand). These problems occupy Chapter IV.

There, it is observed that a number of economists have lately succeeded in devising theoretical analyses which take into account the simultaneous determination of schooling, on-the-job training, hours of work--and sometimes, consumption--over the life cycle of the utility-maximizing individual or household. Models of this sort yield their results in the form of explicit or implicit solutions which describe optimal lifetime trajectories for the variables under the control of the maximizing agent. As one might expect, these solutions, when they can be derived at all, invariably turn out to be complicated nonlinear functions, involving the rate of time preference, the parameters of the static utility function, and other constants having to do with the production and depreciation of human capital. The implied functional forms present numerous difficulties even under the most favourable circumstances, but they are practically impossible to estimate with data sets as large as the one examined here.

Fortunately, it is possible to implement the general notion of simultaneity using a straightforward procedure, which though somewhat lacking in theoretical rigour, may nevertheless prove highly informative. Chapter IV elaborates a two-equation simultaneous system--one linear equation for earnings and one for hours--which appears to capture the essence of the problem. Results, generated by the method of three-stage least squares, are displayed in Chapter V. These may be compared directly with the estimates of Chapter III in order to assess the degree of bias inherent in the single-equation approach. The system estimates, taken on their own, allow one to evaluate the structural parameters which govern the income-hours-schooling interaction.

Readers primarily interested in empirical results are thus referred to Chapters III and V, or to Chapter VI, where the conclusions reached in this dissertation are summarized. Those who wish to review the various theoretical models put forward by the human-capital school may begin with Chapter I.

NOTES

INTRODUCTION

[1]In the field of education and training the most important contributions have been: Gordon Bertram, The Contribution of Education to Economic Growth, Economic Council of Canada, Staff Study No. 12 (Ottawa: Queen's Printer, 1966); Bruce W. Wilkinson, "Present Values of Lifetime Earnings for Different Occupations," Journal of Political Economy, LXXIV (December, 1966), 556-572; Jenny R. Podoluk, Incomes of Canadians (Ottawa: Dominion Bureau of Statistics, 1968), Chapter 5; David A. Dodge, Returns to Investment in Training: The Case of Canadian Accountants, Engineers, and Scientists (Kingston, Ontario: Industrial Relations Centre, Queen's University, 1972); Canada, Statistics Canada, Economic Returns to Education in Canada (Ottawa: Information Canada, 1974).

[2]That significant differences do exist was the finding of Jenny R. Podoluk, "Some Comparisons of the Canadian-U.S. Income Distributions," Review of Income and Wealth, XVI (September, 1970), 279-302, and was earlier hinted in Canada, Economic Council of Canada, Second Annual Review (Ottawa: Queen's Printer, 1965), Chapter 5.

[3]His landmark contributions are: "Investment in Human Capital and the Personal Distribution of Income," Journal of Political Economy, LXVI (August, 1958), 281-302; "On-the-Job Training: Costs Returns, and Some Implications," Journal of Political Economy, LXX (October, Supplement, 1962), 50-79; "The Distribution of Labor Incomes: A Survey," Journal of Economic Literature, VII (March, 1970), 1-28. See also "Education, Experience, and the Distribution of Earnings and Employment: An Overview," in Education, Income and Human Behavior, edited by F. Thomas Juster (New York: McGraw-Hill Book Co., 1975).

[4](New York: National Bureau of Economic Research, 1974).

[5]Dodge, op. cit., relies on a large private survey directed at individuals in a narrow range of high-level occupations. The study issued by Statistics Canada (op. cit.) used microdata drawn from the Labour Force Survey.

[6]Another study based on the Public Use Sample appeared as the present draft was undergoing final editing. See Peter Kuch and Walter Haessel, An Analysis of Earnings in Canada (Ottawa: Statistics Canada, 1979), Catalogue No. 99-758E. An unpublished paper by these authors is cited in the following text.

CHAPTER I

MODELS OF INVESTMENT IN EARNING CAPACITY

Human-capital theorists have emphasized two principal means by which individuals may invest in earning capacity. One is through formal schooling; the other is through training received on the job. In this chapter, we shall consider in turn models that have been designed to account for the income gains associated with each mode of investment. After reviewing these specific elaborations of human-capital theory, we shall examine the broader approach suggested by Ben-Porath. This well-known model admits formal schooling and on-the-job training as special cases within a general framework of income maximization.

At various points in the discussion, we shall turn to existing empirical studies for help in assessing the validity of the human-capital assumptions. We shall not consider in any detail the large body of human-capital research which presupposes the truth of the basic doctrine and seeks only to measure particular parameters, such as the rate of return to education. A selective review of the measurement literature appears in Chapter II.

FORMAL SCHOOLING

The Model

Though simple in appearance, the basic "schooling model" contains all the essentials of the human-capital approach.[1] Individuals who attend school are seen as investing foregone earnings in order to secure additional income during later life. In present-value terms, those who undertake s years of schooling receive[2]

$$V(s) = W(s)\int_s^T e^{-rt}dt = [W(s)/r][e^{-rs} - e^{-rT}]$$

$$= e^{-rs}[W(s)/r][1 - e^{-r(T-s)}] \quad , \ . \ .(1)$$

where T indexes the date of retirement, r stands for some appropriate discount rate, and $W(s)$ signifies the annual wage, assumed constant throughout the individual's working life. Similarly, those who undertake $(s - d)$ years of schooling receive

$$V(s-d) = e^{-r(s-d)}[W(s-d)/r][1-e^{-r(T-s+d)}] \quad .$$

It will be observed that these calculations abstract completely from changes in annual earnings caused by planned or unplanned variations in hours of work.[3]

If we now impose the following condition,

$$V(s-d) = V(s) \quad , \qquad \ldots .(2)$$

and transform the schooling variable so that s-d = 0, we obtain

$$W(s) = W(0)e^{rs} \cdot \frac{1 - e^{-rT}}{1 - e^{-r(T-s)}} \quad ,$$

the fraction on the right-hand side being an adjustment for the finiteness of the working life. If T is large in relation to s, or if T varies in order to make working lives equal whatever the length of schooling,[4] the preceding expression reduces to the simple form

$$W(s) = W(0)e^{rs} \quad ,$$

which may also be written conveniently as

$$\ln W(s) = \ln W(0) + rs \quad . \qquad \ldots .(3)$$

Since $dW(s)/W(s) = r \cdot ds$, we arrive at the conclusion, standard in the human-capital literature, that equal proportionate differences in earnings accompany equal absolute differences in the length of schooling.

An Appraisal

To assess the usefulness of the preceding result for understanding real-world economic behaviour, we must now look carefully at the logic and at the assumptions which underlie it. As a matter of present-value accounting, Equation (1) assumes either that students have no income while attending school or that their earnings just offset tuition and similar direct costs, which are otherwise completely ignored.[5] Furthermore, it is assumed that students derive no consumption benefits from their education, either while attending school or during later years. Nonpecuniary aspects of the jobs associated with different levels of schooling are likewise neglected. The errors thus introduced into the cost-benefit arithmetic may be significant; but as this objection to the model is already well known, there is little need to pursue it here.

More important to the present study is the interpretation of Equation (2). Mincer invokes the condition without comment,[6] though it is crucial to his analysis. One is left to wonder whether it is an identity or a behavioural postulate. If it is an identity, then r must be an ex post internal rate of return; for as the definition requires, r is the discount rate that equates total benefits, given by V(s), and total opportunity costs, given by V(s-d). If r is indeed an ex post rate of return, what economic information does it convey?

Becker has argued[7] that when r exceeds the return on comparably risky investments in physical capital, there is evidence of underinvestment in education. Such reasoning is no doubt correct, but from a policy point of view it is regretably superficial. What we really need to know is *why* the underinvestment occurs. Writers of the human-capital school usually stress the likelihood that imperfections on the supply side of the market restrict the availability of private educational finance. Accordingly, they may favour giving students various subsidies and loans. It may well be, however, that students fail to invest because they perceive barriers to entry on the demand side. Under such circumstances, distributing subsidies will increase educational attainment and, very probably, cause r to fall; but if r falls, it will not be because inefficient shortages of educated manpower are relieved, but rather because graduates spend additional time queuing for preferred employment, or because they crowd into inferior jobs. Unless steps are taken to counteract the demand-side imperfections, further investment in education may involve considerable social waste. This example merely emphasizes the limitations of ex post measurements.

If r is to be interpreted instead as an ex ante rate of return, then Equation (2) must be an equilibrium postulate. As such, it injects into the schooling model a set of implicit hypotheses concerning market behaviour. Although Mincer never really pauses to discuss market processes, it is not very difficult to imagine what a consistent rendering of his model might include.

Elaborating slightly upon Equation (1), we obtain

$$V_i^*(s) = e^{-r_i s}[W_i^*(s)/r_i]\,[1 - e^{-r_i(T-s)}] \quad ,$$

which measures the *ex ante* lifetime earnings of individual i, whose personal discount rate is r_i, and whose wage-rate expectations are summarized by the function $W_i^*(s)$.[8] Let us assume that the individual behaves so as to maximize $V^*(s)$. If circumstances permit an interior maximum, he will then seek to acquire that level of schooling s^* for which $dV^*(s^*)/ds = 0$. The result, omitting a small finiteness correction, is simply

$$\frac{dW_i^*(s^*)/ds}{W_i^*(s^*)} \equiv r_i^* = r_i \quad .$$

Marginal expected returns equal marginal (here average) opportunity cost. Solving this differential equation for s^* yields the desired level of schooling.[9] Notice, however, that the preceding condition is irrelevant unless the graph of the function

$$[dW_i^*(s)/ds]/W_i^*(s) = d \cdot \ln W_i^*(s)/ds$$

intersects r_i from above. In other words, the individual's expected rate of return must decline with s.[10] If not, or if no intersection occurs, the optimal level of schooling will be zero, as high as possible, or indeterminate, depending on the particular circumstance.

Now, to reach the market level of aggregation, we may think of r_i as being drawn from a frequency distribution with mean $\bar{r}$ and variance $Var(r)$. Given information on this distribution, on the distribution of expected wages, and on the process linking expected and observed wages, we can determine, at least in principle, the supply of enrollees as W(s) varies, and ultimately, the total stock of workers at each level of schooling.[11] We thus have a set of long-run supply curves. Presumably, there exists a matching set of demand curves based on the profit-maximizing behaviour of employers.[12] In equilibrium, the curves achieve intersections which enforce an equalization of present values, as Equation (2) requires. The discount rate which makes these present values equal will be that of the marginal investor in formal schooling. The equilibrium structure of wages (earnings) will, finally, be implicit in Equation (3).

By concentrating entirely upon equilibrium positions, Mincer, and Becker as well, avoid the complicated question of disequilibrium adjustment. This tactic achieves great elegance and simplicity, but it leaves in darkness the basic functioning of the labour economy. As Schultz says,

> What we want to know is the relative rates of return to investment opportunities and what determines the change in the pattern of these rates over time. To get on with this analytical task, we must build models that reveal the very inequalities that we now conceal and proceed to an explanation of why they occur and why they persist under particular dynamic conditions.[13]

These "inequalities"--the imperfections and disequilibria which seem to pervade labour markets--have been the concern of many labour economists, especially those writing before the rise of modern human-capital theory;[14] but in the schooling model such disturbances are deemed unimportant.

If the model is to provide anything more than ex post description (however useful that might be for *some* purposes), one must assume that dynamic forces succeed in equating present values, and that they do so, within tolerable limits of approximation, not just "in the long run," but at any moment one might happen to select for empirical study. Without this auxiliary dynamic hypothesis, implementation of the static theory embodied in Equation (3) becomes impossible. Unfortunately, prima facie evidence *against* the equalization assumption is both strong and abundant. Early studies by Houthakker, Hansen, and Hanoch in the United States,[15] and by Wilkinson in Canada[16] show wide variation in the present values of lifetime earnings across schooling groups. Subsequent research in North America and elsewhere has reinforced this finding.[17] One must therefore approach the equalization assumption with some skepticism.

Meanwhile, it is interesting to note that Mincer's preoccupation with equilibrium loci has the effect of suppressing completely the demand side of the labour market. Near the end of *Schooling, Experience, and Earnings* he warns that ". . . the earnings function in this study is a 'reduced form' equation, in which both demand conditions and supply responses determine the levels of investment in human capital, rates of return, and time worked."[18] Yet, no exogenous demand variables actually appear in Equation (3). This supply-side approach to earnings determination contrasts sharply with earlier research. As Bluestone, Murphy, and Stevenson observe:

> Labour market investigation in the 1950's was oriented toward the "demand" side, or industry side, of wage determination. During this period, labour economists concentrated on researching interindustry and interregional wage differentials and developing models to measure the effects of unionization, profits, concentration, and capital intensity on industry rates
>
> . . . .
>
> The 1960's saw a major shift from industry studies to research on human capital. . . . Abstracting from the effect of industry and institutional structure, the human-capital-oriented research focused on the education, skills, training, health, mobility, and attitudes of the labour force. . . . . . .
>
> . . . .
>
> . . . In a "vulgar" or extreme human capitalist approach, all industries are treated as though operating in the same labour market, labour mobility is assumed perfect within skill categories, and because of competition, all industries have the same set of economic and institutional conditions. In this model, all variance in wages, including "equalizing" differences, can be explained by the "supply" characteristics of individual workers.[19]

In view of the strong assumptions needed to guarantee long-run equilibrium, and thereby purge the schooling model of demand-side influences, it would appear wise to consider the weaker, yet more easily defensible analytic notion of short-run or "temporary" equilibrium. In a temporary equilibrium, stocks of human capital--that is to say, the number of workers at each level of schooling--need not "fit" the wage structure implied by Equation (3), given local conditions of demand within regions or industries. Demand conditions then determine the actual wage structure, given the stocks of human capital, which though possibly evolving *toward* long-run equilibrium, are nevertheless fixed in the short run. The result will generally be some departure from long-run equilibrium, which can be explained only by permitting demand-side variables to surface in an expanded reduced-form earnings function.

An expanded model, admitting both demand and supply variables, will be derived and tested in Chapter III. This model may be viewed as an attempt, albeit a crude one, to synthesize the alternative approaches to wage determination discussed by Bluestone, Murphy, and Stevenson.

Supporting Analysis and Extensions

To provide a deeper rationale for the schooling model, Becker has suggested that we view its lone constant r as the outcome of equilibrium, not in the market for labour, but in a set of individual "markets" for human capital.[20] The student-investor, who is the decision-making agent in each market, faces an upward sloping supply of educational finance and a downward sloping demand for educational investment. The supply schedule portrays the marginal interest cost of each dollar committed to schooling, and the demand schedule, the marginal expected yield.[21] By equating these values, the individual maximizes net lifetime earnings. He thus determines the optimal amount to invest in schooling and the equilibrium return on his total investment, much as suggested in the preceding subsection.

This equilibrium return might appear to explain the "r" of Mincer's analysis, except that in Becker's framework the rate in question is a marginal one, based on the *dollar* cost of schooling, whereas, in Mincer's own explicit formulation of the problem it is essentially an average, based on the *time* cost of schooling evaluated at some constant opportunity wage W(0). Mincer's "macro" model, unlike Becker's microeconomic rationale, admits no interim rise in the opportunity wage as schooling progresses, nor does it take into account any possible rise in the interest charges that individuals may have to bear. It treats r as a constant rather than as an equilibrating variable. Any distinction between average and marginal rates of return is therefore unnecessary: the two are the same by assumption. However, as we shall observe in Chapter II, Mincer does not always impose this strong restriction in his empirical work.

It is worth noting that Becker--and Mincer too, for that matter--develop their models without considering the rate of time preference. They focus upon the maximization of earnings, not utility. Thus Becker, most paradoxically, mimics the neoclassical theory of investment in physical capital by assuming, implicitly, that consumption and investment in human capital can be made analytically independent. The individual undertakes whatever investment is needed to maximize earnings, and then, treating maximized earnings as a constraint, spreads consumption optimally over his life cycle in accordance with the market rate of interest and his rate of time preference.[22]

The trouble with this approach in Becker's case is that it requires the market for consumption loans to be isolated, somewhat implausibly, from the market for investment finance. Otherwise, the amount an individual borrows for the purpose of consumption spreading will affect the terms under which he may borrow for the purpose of investment. An individual who is an efficient maximizer (of utility) will therefore plan his consumption and investment simultaneously. Perfect loan markets, with perfect arbitrage between them, would restore independence; but Becker has assumed the contrary. As we shall see in Chapter IV, models based on utility maximization are capable of handling such an assumption in principle, although they typically shy away from the very great complexities involved.[23]

The chief use of Becker's model, flawed or not, has been to analyze cross-sectional relationships between rates of return and the level of schooling. For Becker and fellow human capitalists, the demand curves of the model measure individual ability, and the supply curves, opportunity. If the variance of ability within the population exceeds the variance of opportunity, the resulting scatter of individual equilibria will tend to describe a positively sloping line; the more volatile demand curve will "identify" the supply schedule. We shall then observe a positive association between schooling and the rate of return. In the reverse case, we shall witness a negative association, and in the case of equal variances, no correlation whatever. The model is thus capable of accommodating *any* empirical outcome.

In light of the remarks already directed toward Mincer's version of the schooling model, it should come as no surprise to find Becker interpreting the demand side of his own analysis solely as a means of portraying the personal characteristics of individuals. Though Becker deals only with "ability" (a composite of various personal attributes), the demand curves which he postulates must surely depend not only upon this factor but also upon (the individual's perception of) general labour-market conditions. Nevertheless, individuals of equal ability always face identical demand curves. "Inequality of opportunity" cannot occur through unequal access to high-paying jobs in favoured regions or industries, but only through unequal access to investment finance.[24]

In an interesting attempt to apply Becker's demand-and-supply framework, Haessel and Kuch[25] postulate an explicit reduced-form equation for r_i, namely,

$$r_i = a_0 + \sum_{k=1}^{K} a_k A_{ik} , \qquad \ldots .(4)$$

where the a's are reduced-form coefficients, and the A's stand for personal attribute variables.[26] Substituting (4) into (3) yields

$$\begin{aligned} \ln W_i(s) &= \ln W_i(0) + (a_0 + \sum_{k=1}^{K} a_k A_{ik}) s_i \\ &= \ln W_i(0) \quad a_0 s_i + \sum_{k=1}^{K} a_k (s_i A_{ik}) . \qquad \ldots .(5) \end{aligned}$$

Given the form of the K additional variables $X_{ik} = s_i A_{ik}$ appended to the basic equation, one might label (5) the "interactions model."

On ad hoc grounds for the most part, Haessel and Kuch select seven characteristics--religion, ethnicity, occupation, class of worker (salaried or self-employed), period of immigration, marital status, and place of schooling--to define the A_{ik}. In so doing, they explore a number of worthwhile hypotheses, but they do not exhaust the possibilities of the model, given the available data. In particular, the authors do not consider the effects that region and industry of employment might have on the rate of return, as measured in the short run or under conditions of sustained market imperfection. Hypotheses pertaining to these factors will be tested, within an interactions framework, in Chapter III.

Although we have so far dealt with the schooling model, strictly speaking, as a theory of earnings *determination,* it has actually been applied in its purest and simplest form as a theory of earnings *distribution.*[27] Observe that if we take variances on both sides of (3) and assume W(0) to be independent of r and s,[28] the general result is

$$\begin{aligned}\mathrm{Var}(\ln W) &= \mathrm{Var}[\ln W(0)] + \mathrm{Var}(rs)\\ &= \mathrm{Var}[\ln W(0)] + \bar{r}^2 \cdot \mathrm{Var}(s) + \bar{s}^2 \cdot \mathrm{Var}(r)\\ &\qquad + 2\bar{r}\bar{s}\,\mathrm{Cov}(r,s) + R(r,s), \quad \ldots(6)\end{aligned}$$

where R(r,s) is a function involving certain expected values and Cov(r,s).[29] However, if r and s are also independent of one another, (6) reduces to

$$\begin{aligned}\mathrm{Var}(\ln W) &= \mathrm{Var}[\ln W(0)] + \bar{r}^2 \cdot \mathrm{Var}(s) + \bar{s}^2 \cdot \mathrm{Var}(r)\\ &\qquad + \mathrm{Var}(r) \cdot \mathrm{Var}(s). \quad \ldots(7)\end{aligned}$$

In both cases, the left-hand side turns out to be an already familiar measure of earnings inequality; hence, the distributional implications of the model appear immediate and direct. One should of course remember that Var(ln W) is by no means the only plausible measure of inequality, and that its adoption for policy purposes must ultimately rest upon normative considerations.[30]

Writers of the human-capital school--Becker, Chiswick, and Mincer--adhere consistently to the assumption that r and s behave as independent random variables, and so are content to apply (7) in attempting to analyse distributional questions. They obtain the unambiguous result that inequality depends in positive fashion upon the means and the variances of r and s. This prediction with respect to $\bar{s}$ is somewhat surprising, in view of the levelling effect popularly credited to education. One must bear in mind, however, that policies designed to raise $\bar{s}$ will seldom leave Var(s) unchanged; it is unlikely, in other words, that all groups will receive equal increments of schooling. The practical outcome will depend on who gets the additional education. Furthermore, it is difficult to think that $\bar{r}$ would remain constant in the face of an increase in $\bar{s}$. Ceteris paribus arguments based on (7) may thus prove misleading.

As we have seen, the independence assumption, which ultimately supports the preceding results, implies in the context of Becker's analysis that the dispersion of "abilities" and the dispersion of "opportunities" throughout the population must be roughly equal. Mincer contends: "There are no a priori reasons for specifying which dispersion is greater, and the empirical evidence suggests there is little if any correlation between rates of return and quantities invested across individuals."[31] As a matter of fact, evidence for the United States of a significant relationship between r and s is rather widespread. The work of Hansen and of Hanoch,[32] and Mincer's own findings,[33] taken at face value, reveal an apparent negative association, but Mincer dismisses these results as the effect of not holding hours of labour constant.[34] We shall examine this argument carefully in Chapter II and test it by alternative methods in Chapters III and V. For the time being, it is sufficient to note that what seems true of the United States may not be true of Canada.

If years of schooling and the rate of return are, in fact, negatively correlated, then (6) rather than (7) is the appropriate formula. Since by hypothesis $\mathrm{Cov}(r,s) < 0$, the relationship between $\mathrm{Var}(\ln W)$ and $\bar{s}$ is no longer unambiguously positive: an increase in the general level of education need not generate an increase in inequality. Using Hanoch's rate-of-return estimates, Marin and Psacharopoulos produce simulations which do exhibit a decline in inequality as the result of such an increase.[35] The popular view of education thus receives some comfort.

When we come to consider the entire distribution of earnings rather than merely its variance, inspection of (3) is enough to show[36] that if schooling is normally distributed, the distribituon of earnings will be lognormal, or more significantly, that the distribution of earnings will *not* be lognormal (as is sometimes supposed) *unless* schooling is normally distributed. In general, the distribution of earnings will be skewed to the right--a customary finding--as long as the distribution of schooling is not radically skewed to the left.

Oulton, in particular, finds this yield of theoretical predictions unimpressive.[37] The problem, he says, is that the human-capital approach to distribution theory is incomplete: "The distribution of income is made to depend on the distribution of education (or training in general), but the latter is unexplained."[38] Proceeding out of skepticism, Oulton looks for the end of the analytical chain in the area of marginal productivity theory. He postulates an aggregate CES production function

$$Q = (a_0 L_0^{-b} + a_1 L_1^{-b} + \cdots + a_n L_n^{-b})^{-1/b} , \qquad \ldots .(8)$$

which makes distinct inputs--that is to say, imperfect substitutes--of workers who differ by level of education.[39] Here, Q stands for real output, and L_s for the number of workers with s years of schooling ($s = 0, 1, \cdots, n$); the a_s reflect such workers' "inherent productivity"; and below, $\sigma = 1/(1+b)$ will be used to denote the constant elasticity of substitution. Physical capital is ignored.

If workers are paid their marginal products, it is easy to show that

$$W_s = W_0(a_s/a_0)(L_s/L_0)^{-1/\sigma} . \qquad \ldots .(9)$$

Substituting (9) into (3) and solving reveals

$$L_s = L_0(a_s/a_0)^{\sigma} e^{-rs\sigma} \quad . \qquad \ldots .(10)$$

Finally, if we assume for expositional convenience that (a_s/a_0) takes the form $e^{\gamma s}$, where γ is possibly a function of s, Equation (10) becomes

$$L_s = L_0 e^{(\gamma - r)s\sigma} \quad . \qquad \ldots .(11)$$

This expression implies the form of the schooling distribution. If the latter is to display the humped character required by (3) to explain the observed distribution of earnings, inspection of Equation (11) suggests that γ must first exceed and then fall below r as s rises. In other words, the a_s must conform to a particular pattern. Oulton concludes that

> . . . there are no a priori reasons for expecting this particular pattern of 'inherent productivity' to be found in the real world. If, therefore, the model is thought to be an adequate description of reality, it would be for essentially accidental reasons. . . .[40]

Owing to the somewhat restrictive nature of the production specification advanced in (8), it is perhaps a little unwise to accept this statement without further analysis. One might at least consider the possibility that, in the long run, technology may be endogenous. If the a_s eventually adjust to accommodate a schooling distribution determined, say, by ability or socio-economic background, the resulting pattern of coefficients will be far from "accidental." To confirm this speculation here, within a rigourous maximizing framework, would unfortunately require a major disgression. Therefore, let us simply accept Oulton's essential point: that in the short run, most certainly, and perhaps also in the long run, human-capital theory is suspect because it ignores the demand side of the earnings-distribution problem.

ON-THE-JOB TRAINING

Mincer's Theory

The schooling model we have just examined actually arises as a special case within the more general framework offered by human-capital theorists to account for on-the-job training and other forms of postschool investment. Mincer's current approach to on-the-job training is a straightforward elaboration of the model suggested originally by Becker and Chiswick.[41] This treatment rests on the distinction between an individual's actual earnings after p years of work experience, $W_i(p)$, and his earning capacity, $E_i(p)$. The latter equals $W_i(p)$ plus $C_i(p)$, the income foregone in order to attain further skills or earning capacity.

If we now think of each increment of foregone earnings as yielding some rate of return r_p, we may write (in discrete form) the accounting identity

$$E_p \equiv E_0 + \sum_{t=0}^{p-1} r_t C_t \equiv W_p + C_p \qquad \ldots .(12)$$

where the subscript relating to individuals has been dropped for convenience. The next step is to make investment C_p a function of earning capacity; that is,

$$C_p = k_p E_p \ , \qquad 0 \leqq k_p \leqq 1 \ .$$

One may interpret k_p as the proportion of total "market time" devoted to skill acquisition during year p. The logic of (12) then implies

$$E_p = E_{p-1} + r_{p-1} C_{p-1} = E_{p-1}(1 + r_{p-1} k_{p-1}) \ .$$

By successive substitution, we obtain

$$E_p = E_0 \sum_{t=0}^{p-1} (1 + r_t k_t) \quad ,$$

which is approximately equivalent to

$$\ln E_p = \ln E_0 + \sum_{t=0}^{p-1} r_t k_t \quad , \qquad \ldots .(14)$$

as long as $r_t k_t$ is small. Since $E_p = W_p/(1 - k_p)$, we finally arrive at

$$\ln W_p = \ln E_0 + \sum_{t=0}^{p-1} r_t k_t + \ln (1 - k_p) \ . \qquad \ldots .(15)$$

During formal schooling, individuals may be thought to specialize in the production of human capital, and thus for $t = 0, 1, \cdots, s$, $k_t = 1$. In this case, if the rate of return is the same in each period, (15) reduces to (3), the basic schooling model, with E_0 redefined to mean earning capacity in the absence of both education *and* experience (that is $E_0 = W_0$).

Allowing separate, though constant rates of return (denoted here by r^e and r^x, respectively) to each of these investment modes, Mincer partitions (15) in the necessary manner to obtain

$$\ln W_p = \ln W_0 + r^e s + r^x \sum_{t=0}^{p-1} k_t + \ln (1 - k_p) \quad . \ . \ . \ . \ .(16)$$

However, if k_t declines monotonically over the individual's working life (as will be discussed later), this model implies that measured earnings W_p rise steadily until retirement. To explain the slight "hump" sometimes detected in age-earnings profiles, one must introduce the concept of depreciation.[42] If human capital depreciates at some constant rate d, then

$$E_p = E_{p-1} + r_{p-1} C_{p-1} - dE_{p-1} \quad ,$$

which leads eventually to

$$\ln W_p = \ln W_0 + (r'^e - d)\, s + r^x \sum_{t=0}^{p-1} (k'_t - d/r^x) + \ln (1 - k'_p) \quad .$$

$$. \ . \ . \ . \ .(16')$$

One may think of $r^e = (r'^e - d) \gtreqless 0$ as the net rate of return to schooling and of $k_p = (k'_p - d/r^x)$ as the net propensity to invest in human capital. Primes denote the corresponding gross values. Because elements of the summation on the right-hand side of (16') may turn out to be negative, it is now possible for W_p to decline over some interval--presumably near the end of the individual's working life, when he is unable to amortize large gross investments. Whatever the precise empirical result, Equation (16') stands as the culmination of Mincer's theoretical analysis: it is the model for which he attempts to derive an operational likeness.

By recognizing opportunities for postschool investment, Mincer and his fellow human capitalists provide a convenient rationale for the observed tendency of individual earnings to rise over most (if not all) of the life cycle. Moreover, as long as k_p decreases with time, the expanded model implies that earnings profiles, even in the absence of variations in labour supply, will appear concave from below. The model thus "explains" one of the stylized facts connected with life-cycle earnings.

The final important implication of Mincer's analysis has to do with his controversial notion of "overtaking." Because postschool investors sacrifice potential income, they at first earn less than hypothetical noninvestors, whose earnings profiles are assumed to remain horizontal. Later, as returns accrue and as commitments of potential income decline, investors earn more. If we focus momentarily upon dollar costs and returns, then, at the overtaking year of experience $\tilde{p}$,

$$W_{\tilde{p}} = W_s + r^x \sum_{t=0}^{\tilde{p}-1} C_t - C_{\tilde{p}} = W_s , \qquad \ldots .(17)$$

if and only if

$$r^x \sum_{t=0}^{\tilde{p}-1} C_t = C_{\tilde{p}} .$$

If annual dollar investments were constant at $\bar{C}$ during the first p years after school leaving, we should obtain $r^x\tilde{p}\bar{C}$, which means that $\tilde{p} = 1/r^x$. On the other hand, if dollar investments decline as we expect, it is easy to show that $\tilde{p} < 1/r^x$. Hence, we can place an upper bound on $\tilde{p}$, provided we know r^x. Mincer assumes that r^x ". . . is not very different from the rate of return as usually calculated [for education] . . . ," thus making $\tilde{p}$ "a decade or less."[43]

Now, if rates of return and the detailed pattern of investment, as opposed to the *total* planned accumulation, do not vary inordinately across individuals, overtaking will occur in practice within a relatively narrow band of years after school leaving. In other words, the earnings profiles of large and small postschool investors, and of noninvestors, if there are any, will be observed to intersect at roughly the same point. The experience cohort thus identified should exhibit less inequality than others in the working population, although strictly speaking, such an inference depends on the further assumption that there exists an appropriately small correlation between potential earnings at school leaving and the propensity to engage in postschool investment.[44] The cross-cohort patterns of inequality found by Mincer actually display the expected minima only in the case of high-school graduates, leading him to conclude in the contrary instances that the correlation just named must *not* be sufficiently small.[45] Thus, again, the human-capital approach proves capable of accommodating any conceivable result.

An Appraisal

One may surely be forgiven for remarking that just a single unambiguous prediction--that earnings profiles are concave--does not seem a very substantial dividend with which to repay the preceding analysis. Consistency with stylized fact is comforting but inconclusive, particularly in the face of competing explanations. One of these holds that concave earnings profiles are largely the result of biological factors connected with aging. If this hypothesis is true, age should figure at least as prominently as experience in the determination of cross-sectional earnings. The rare data sets which supply information on both of these independent variables unfortunately generate mixed qualitative results, although the weight of quantitative evidence seems to rule out extreme versions of the age hypothesis. Malkiel and Malkiel find that age is not significant when included in a regression along with experience.[46] However, studies of the engineering profession, by Cain, Freeman, and Hansen, and by Klevmarken and Quigley, uncover a small but not unimportant effect of age on earnings.[47] Lazear encounters a relatively strong age effect,[48] and

Psacharopoulos, observing a backward economy, reports that even illiterate, unskilled workers exhibit concave earnings profiles.[49] One must conclude that investment behaviour, represented empirically by years of work experience, may not be the *sole* determinant of concavity.

A stronger objection to the postschool investment model arises from the potential significance of costless learning by doing. As Blaug observes, ". . . any psychological theory of 'learning curves,' in which appreciation over time is partly but only partly offset by depreciation and obsolescence, will likewise account for concave age-earnings profiles."[50] If learning by doing predominates over forms of training which use real resources or sacrifice output, the investment interpretation of earnings profiles appears to lose much of its appeal, since an activity which is costless and as inexorable as the passage of time cannot be the subject of an investment decision.

However, in *Human Capital,* Becker argued that labour mobility and competition for jobs would effectively eradicate costless opportunities for learning.[51] If such opportunities ever arose, workers would crowd into them, forcing wage rates to adjust until productivity-constant and productivity-enhancing employment yielded the same present value of lifetime earnings. In equilibrium, the rising income profiles again intersect the horizontal ones, and workers must make a choice. As in the case of the schooling model, the human-capital interpretation of on-the-job training depends completely on the belief that competition succeeds in equating present values.

Whether competitive forces in real-world labour markets actually possess such power is clearly open to debate.

In general, the objections raised against the schooling model seem to apply with equal force to the expanded theory. If anything, market processes and the role of demand appear more deeply submerged in the latter than in the former. Equations (12)-(16') might very easily be regarded as identities with no direct behavioural significance. The model contains, in a sense, too many "degrees of freedom"; because potential income is unobservable, so is the crucial investment parameter k_p. Though, as we shall see in the next section, the income maximization models put forward by some human-capital writers do make one or two predictions concerning the time path of k_p, the restrictions placed, by inference, upon observable quantities like measured income are normally too weak to generate a very powerful or discriminating test of the theory.

Supporting Arguments

To the extent that human capitalists concern themselves at all with market functioning and firm behaviour, it is usually in order to explain the mechanism through which workers undertake investment "expenditures" while on the job. That full-time workers, like full-time students, forego income, and do so to a planned degree (given by k_p), may not be immediately obvious. In the case of foregone income invested in generally marketable skills, Becker's well-known response[52] was to argue that because a trained worker could always obtain his marginal product in a competitive labour market, that worker would receive the entire return on any investment made by him, and would, if necessary, be willing to pay its full cost. An employer who had to bear the cost initially but who could guarantee himself none of the return (because the worker might quit) would require compensation for any training provided. Untrained workers pay the needed compensation by accepting a wage which falls short of their marginal product.

In the model ingeniously devised by Rosen,[53] such workers choose the amount of their investment by selecting a job with the appropriate characteristics. Rosen states:

> The nature of the market is such that workers have their choice among all-or-nothing bargains or 'package deals,' in which they simultaneously sell the services of their skills and 'purchase' a job offering a fixed opportunity to learn. By the same token firms purchase services of skills and at the same time 'sell' jobs offering learning possibilities. The labor market provides a broad range of choice in these matters. . . .
> . . . Prices of jobs could be either explicit or implicit, but the distinction is of no analytical importance. . . . Ordinarily, investment costs are simply subtracted from gross pay and no explicit price need be quoted.[54]

In Rosen's model it makes no difference whether firms supply costly forms of training or costless learning by doing. Both in the market for existing skills and in the market for skill development, competition assures a simple, determinate result. Firms offer a profit-maximizing menu of learning opportunities, and over the life cycle, workers move from job to job (varying k_p) in pursuit of their investment goals.

It must be conceded that this view of on-the-job training and life-cycle investment places a rather heavy information burden upon both parties to the learn-and-earn bargain. Workers and employers must be able to predict, within tolerable limits, the training characteristics of a great many jobs. Whether they can do so with sufficient accuracy to make the theory realistic is a difficult question. Furthermore, it might appear to some that the notion of workers' having to change jobs continually in order to fulfill their investment plans seriously misrepresents the nature of occupational mobility in the labour market. As Blaug says skeptically, ". . . it is . . . doubtful that all interoccupational, and even more intraoccupational, movements of labor can be reduced to the action of sowing and reaping the advantages of labor training. . . ."[55] That workers remain in essentially the same occupation and "ride" a fixed learning curve seems, all in all, a simpler explanation for what we observe in the labour market.

In the case of training which is valuable only to the firm which provides it, Becker's argument[56] was that employers could collect the entire return and would therefore be willing to pay the entire cost, but that they would more likely *share* costs and returns with workers in order to discourage turnover. By promising workers a rising experience profile of wage rates, employers could reduce quits and, hence, the loss of investment in "specific training." The wage profile which kept such losses to a minimum would implicitly determine the equilibrium sharing of costs and returns.

In a recent article, however, Donaldson and Eaton[57] contend that the idea of shared investment is mistaken. According to their definition, "sharing" occurs only if the wage profile offered to the worker makes him better off in present-value terms than he would be in alternative employment. It is immediately obvious that competition among workers will never permit sharing in this sense. Superior opportunities will always be eroded. The firm will manipulate the wage profile in order to minimize the loss of experienced workers; but since its wage bill (in present-value terms) is fixed, it must ultimately collect the total net benefit of any specific training it decides to undertake. Granting the important point with regard to sharing, one should not however be misled by the Donaldson-Eaton analysis into thinking that specific training does not pose an investment problem *from the worker's viewpoint*. When offered a rising wage profile, as opposed to a flat one in alternative employment, the worker must still decide which to accept; and for this purpose he must perform an investment calculation. The Donaldson-Eaton analysis, although sufficient to make its point, suffers to a certain extent from its failure to elaborate the worker's decision problem.

One may also question whether it is appropriate to assume competitive behaviour in modelling the relationship between firms and their employees. As Reder commented in his review of Human Capital,

> . . . an individual employee can, by quitting, impose a loss on an employer of his (the employer's) whole share of the return on training. Hence, any share of the return that a worker lets an employer keep makes that employer better off than he

> would have been if the worker had quit. On the other hand, it is obvious that by discharge, the employer can impose an analogous loss on the worker.
>
> Thus is generated the zone within which bargaining power, strategic skill, institutional rules, etc., determine wage rates.[58]

However, if workers (and firms) accurately foresee these bargaining possibilities, the gains or losses which flow from them will presumably affect the initial decision of whether or not to accept employment (or hire) at a given starting wage. Competition for *opportunities* to bargain should negate any advantages or disadvantages which bargaining might otherwise entail.

As far as Mincer is concerned, the analytical differences between general and specific training are of little ultimate consequence, since their separate influences upon age-earnings profiles are empirically indistinguishable, given the available data. Both imply, very simply, that earnings (exclusive of depreciation) rise with work experience. In the absence of detailed information on learning curves and on the direct and indirect expenditures of firms, experience must serve as a proxy for all the various modes of on-the-job training. In fact, as we shall see in the next chapter, experience can itself be estimated from census data only by means of a further proxy.

GENERAL THEORIES OF INCOME MAXIMIZATION

In the Becker-Chiswick-Mincer analysis, individuals decide upon the amount and timing of their investment in human capital by choosing a sequence of values for k'_p. If the foregoing model is to be understood as something more than a tautology in which $k'_p = C_p/E_p$ ex post, one must supply a behavioural theory to predict the course of this variable over the individual's life cycle. The first to approach the task was Ben-Porath.[59] His model, and the extension provided by Haley,[60] may be termed "general" insofar as they treat schooling and on-the-job training as special cases within a choice-theoretic framework. That framework is nevertheless one of income rather than utility maximization. In the present context both yield the same result, since the authors continue to assume a single good, ignoring leisure.

Ben-Porath's essential contribution to the analysis was the idea of an individual production function for human capital. Applying this device, one assumes that the individual "manufactures" increments Q_H of human capital by bringing together purchased inputs D and a portion of some existing capital stock H. The production function, in its most general form, may be written

$$Q_H = F(k', H, D) \quad . \qquad \ldots .(18)$$

However, Ben-Porath invokes the so-called "neutrality assumption" to obtain

$$Q_H = f(k'H, D) \quad . \qquad \ldots . (18')$$

Here, human capital is treated as an augmenting factor, and k'H represents effective investment time. If this time were sold in the labour market, it would bring earnings of w(k'H), where w signifies the fixed rental price of human capital. "Neutrality" hinges on the assumption that effective investment time and effective work time incorporate the same augmenting factor, H. Thus, human capital increases earning potential and the ability to generate further earning potential in exactly the same proportion.

In Haley's somewhat simplified version of the model, purchased inputs disappear, and the production function becomes

$$Q_H(t) = \alpha I(t)^{\mu} \quad , \qquad \ldots . (19)$$

where

$$I(t) = k'(t)\, H(t) \quad .$$

All variables are treated as continuous functions of time. The first parameter, α, measures individual efficiency in human-capital production, and the second μ, denotes the level of returns to scale. Unless returns to scale are declining $(0 < \mu < 1)$, the model will not yield an acceptable solution.

In view of depreciation, the individual's stock of human capital must evolve according to the differential equation

$$\dot{H}(t) = Q_H(t) - dH(t) \quad . \qquad \ldots .(20)$$

Earning capacity is simply $E(t) = wH(t)$, and "disposable earnings" are given by

$$W(t) = wH(t) - wI(t) = [1 - k'(t)]\, wH(t) \quad . \qquad \ldots .(21)$$

The problem for the individual is to choose $k'(t)$ in order to maximize

$$J = \int_0^T W(t)e^{-rt}\,dt \quad , \qquad \ldots .(22)$$

subject to (20) and (21). Together with the boundary restrictions

$$H(t) \geqq 0 \quad , \quad I(t) \geqq 0 \quad , \quad H(t) - I(t) \geqq 0 \quad ,$$

and some initial condition $H(0) = H_0$, Equations (20)-(22) define a relatively simple problem in control theory.

As usual, the solution procedure generates a set (more specifically, a continuum) of shadow prices for human capital, namely:

$$\lambda(t) = [w/(r+d)]\,[1 - e^{(r+d)(t-T)}] \quad , \quad 0 \leqq t \leqq T \quad .$$

These decline over the life cycle because of the dwindling opportunity to amortize new investment prior to the fixed retirement date. The reasonable supposition that the stock of human capital becomes worthless at retirement justifies the transversality condition

$$\lambda(T)\, H(T) \;=\; 0 \quad . \qquad \ldots .(25)$$

Wherever the individual attains an interior solution, he optimizes by choosing $k'(t)$, and hence $Q_H(t)$, so that the marginal cost of producing the desired amount of human capital equals the ruling shadow price, $\lambda(t)$. Since $\lambda(t)$ falls continuously over time, and since marginal cost is perforce assumed to be a rising function of human-capital output,[61] the increments $Q_H(t)$ added to the human-capital stock must decline monotonically over the life cycle. Effective investment time $I(t)$ must also decline monotonically; to be specific,[62]

$$\dot{I}(t) \;=\; I(t) \cdot \frac{-(r+d)}{1-\mu} \cdot \frac{e^{(r+d)(t-T)}}{1-e^{(r+d)(r-T)}} \;<\; 0 \qquad \ldots .(26)$$

The behaviour of $k'(t)$ is more difficult to establish. From the definition $I(t) = k'(t)H(t)$, and from Equation (20), one may deduce that

$$\dot{k}'(t)/k'(t) \;=\; \dot{I}(t)/I(t) - \dot{H}(t)/H(t) \quad , \qquad \ldots .(27)$$

or

$$\dot{k}'(t) \;=\; k'(t)[\dot{I}(t)/I(t) - Q_H(t)/H(t) + d] \quad .$$

The sign of the bracketed expression appears indeterminate, unless $d = 0$. Then, without question, $\dot{k}'(t) < 0$. In general, it would seem that fulfillment of the optimal plan might require $k'(t)$ to increase over some interval late in the individual's life cycle, when $\dot{H}(t)/H(t) < 0$. However, this conclusion cannot be accepted without first substituting, for the endogenous variables in (27), their equivalents in terms of the model parameters, r, d, α, μ, T, and H_0. The resulting expression for $k'(t)$ is virtually impossible to deal with analytically. Instead, $k'(t)$ was simulated numerically for a wide range of parameter combinations. In every case, $k'(t)$ declined monotonically. The simultations also confirm Haley's assertion that $k'(t)$ must display an inflection point. Results verify that the function declines first at a decreasing, and later at an increasing rate.[63] At retirement, of course,

$$k'(t) = k'(T) = 0 \quad .$$

At the opposite end of the age scale, the foregoing analysis may not apply, for individuals typically appear not to achieve interior maxima. When $\lambda(t)$ is high because of the long amortization period in prospect at the beginning of the economic life cycle, optimization according to the rule $MC(t) = \lambda(t)$ may require the investment of more human capital than the individual currently owns. At such times, the boundary condition $H(t) - I(t) \geqq 0$ holds with equality, and the individual specializes in the production of human capital, setting $k'(t)$ equal to one. Though it is natural to identify the period of specialization with that of formal schooling, the two need not be coextensive. Specialization may very well cease before schooling finishes; indeed, many "full-time" students devote a considerable number of hours to market work. Such behaviour is consistent with the theory, since the optimal plan may dictate $k'(t) < 1$ for some $t < s$.

The length of the specialization period, whether or not it falls short of s, is determined endogenously as part of the optimization programme. Haley shows[64] that the length depends positively upon α, the individual's personal efficiency parameter, and negatively upon r, d, and H_0. The latter is of course the individual's initial endowment of human capital. That α an H_0, which may be positively correlated, should have opposite effects on the period of specialization is a particularly intriguing outcome of the analysis.

Unfortunately, the broad implications of the model stand up rather poorly in the face of existing evidence. A second-derivative test conducted by Ben-Porath[65] makes use of the fact that

$$\frac{\partial(\dot{I}/I)/\partial t}{(\dot{I}/I)} = \frac{r+d}{1-e^{(r+d)(t-T)}} > 0 \quad . \qquad \ldots .(28)$$

This equation predicts "the rate at which the decline in investment over the life cycle should accelerate."[66] Employing the data from Mincer's 1962 study of on-the-job training, Ben-Porath finds that investment (inferred from age-earnings profiles) falls much more rapidly than one would expect on the basis of Equation (28). Moreover, estimates of μ, obtained by combining (28) and (26), suggest that returns to scale are nearly constant ($\mu \cong 1.0$). This result tends to contradict the crucial assumption upon which the model rests.

One explanation may be that the neutrality hypothesis is false.[67] If human capital is biased towards the market, and if the bias increases with time, investment will in fact decline more rapidly than Equation (26) predicts. Whether the decline will *accelerate* nevertheless appears uncertain.[68] Still, there does not seem to be any weaker or more general hypothesis which preserves testability. One cannot use an equation like (28), for example, to identify a further set of bias parameters. On the other hand, if the only conceivable structure one may impose upon the model--the neutrality hypothesis--is rejected by the evidence, the chief advantage of Ben-Porath's explicit maximization approach disappears. One might just as well employ the simpler, ad hoc analysis put forward by Mincer.

Other problems may of course account for the apparent failure of the Ben-Porath model. Three that have been discussed in the literature are: vintage effects that may distort cross-section age-earnings profiles;[69] life-cycle variation in hours of work;[70] and the use of contradictory assumptions in the construction of investment profiles.[71] Brown[72] proposes remedies for all three, but his results are not wholly encouraging. Though he obtains plausible estimates of μ, the values implied for r appear unreasonably low.

In another study, Heckman[73] once again encounters constant returns to scale. Upon estimating k'(t), he finds an initial segment of the function that is positively sloped, and second-order properties that are the reverse of those forecast by Haley. On the other hand, Haley's own research,[74] using grouped data and a complicated non-linear estimation procedure, strongly supports the Ben-Porath theory. Parameter estimates fall within reasonable limits and display relatively small variances. One is therefore left with an indecisive result and a need for further, detailed research.

APPENDIX I

THE EFFECT OF MARKET BIAS ON THE OPTIMAL INVESTMENT PROFILE

We have seen in the foregoing text that if neutrality holds, it is possible to entertain a human-capital production function of the form

$$Q_H = \alpha(k'H)^{\mu} = \alpha I^{\mu} \quad .$$

Marginal cost is thus given by

$$MC = w/(\partial Q_H/\partial I) = (w/\alpha\mu)\, I^{1-\mu} \quad .$$

Optimization according to the rule MC = λ implies that

$$(w/\alpha\mu) I^{1-\mu} = [w/(r+d)][1 - e^{(r+d)(t-T)}]$$

$$\therefore \quad I = \left\{ \frac{\alpha\mu}{r+d} [1 - e^{(r+d)(t-T)}] \right\}^{1/(1-\mu)} \quad . \qquad \ldots\ldots(\text{A.I.1})$$

Now, to insert the notion of market bias, we may rewrite the production function in the following manner:

$$Q_H = \alpha(bI)^{\mu} = \gamma I^{\mu} \quad , \qquad \ldots\ldots(\text{A.I.2})$$

where

$$\gamma = \alpha b^{\mu}$$

and

$$b = b(t) \quad .$$

If b, the bias parameter, equals one, we have neutrality. If $0 < b < 1$, human capital is biased towards the market: the current increment adds Q_M to earning capacity but only bQ_H to potential investment input. If $b > 1$, human capital has an "investment bias." We may suppose that b is an exogenous function of time (age).

It should be obvious from the preceding derivation that

$$I' = \left\{ \frac{\gamma\mu}{r + d} [1 - e^{(r+d)(t-T)}] \right\}^{1/(1-\mu)} < I \quad , \quad \text{if} \quad b < 1 \quad . \qquad \ldots . (A.I.3)$$

At all points during the nonspecialization phase of the life cycle, market bias reduces the level of investment in human capital. Market bias also reduces the length of the specialization phase. Both effects are due to the increase in marginal cost.

Differentiating (A.I.3) in logarithmic form yields

$$\frac{\dot{I}'}{I'} = \frac{\dot{\gamma}}{(1 - \mu)\gamma} + \frac{-(r + d)e^{(r+d)(t-T)}}{(1 - \mu)[1 - e^{(r+d)(t-T)}]} \quad , \qquad \ldots . (A.I.4)$$

which is unambiguously negative if $\dot{\gamma} < 0$ --that is, if market bias increases with age. One might reasonably expect this condition to hold. If so, comparison of (A.I.4) and (26) demonstrates that $|\dot{I}'/I'| > |\dot{I}/I|$. Market bias causes investment to decline more rapidly (in proportional terms) than under conditions of neutrality. However, if market bias is constant ($\dot{\gamma} = 0$), $\dot{I}'/I' = \dot{I}/I$; and the rate of decline is unaffected.

For convenience in what follows, let us now implicitly define some new notation by re-expressing (A.I.4) as

$$\frac{\dot{I}'}{I'} = \frac{\dot{\gamma}}{z\,\gamma} + \frac{-RX}{z(1 - X)} \quad .$$

Differentiating once more, we obtain

$$\frac{d(\dot{I}'/I')}{dt} = \frac{\gamma\ddot{\gamma} - \dot{\gamma}^2}{z\,\gamma^2} + \frac{-R^2X}{z(1 - X)^2}$$

$$= \frac{(\gamma\ddot{\gamma} - \dot{\gamma}^2)(1 - X)^2 - \gamma^2R^2X}{z\gamma^2(1 - X)^2} \quad ,$$

where $\ddot{\gamma} = d^2\gamma/dt^2$. We wish to divide the preceding expression by

$$\frac{\dot{I}'}{I'} = \frac{\dot{\gamma}(1 - X) - \gamma RX}{z\gamma(1 - X)} \quad .$$

The result is

$$\frac{d(\dot{I}'/I')/dt}{\dot{I}'/I'} = \frac{(\gamma\ddot{\gamma} - \dot{\gamma}^2)(1 - X)^2 - \gamma^2 R^2 X}{\gamma(1 - X)[\dot{\gamma}(1 - X) - \gamma R X]} \quad . \qquad \text{. . . . .(A.I.5)}$$

We must finally compare (A.I.5) and (28). In our present notation the latter is simply $R/(1 - X)$. Market bias will increase the relative rate of deceleration if

$$\frac{(\gamma\ddot{\gamma} - \dot{\gamma}^2)(1 - X)^2 - \gamma^2 R^2 X}{\gamma(1 - X)[\dot{\gamma}(1 - X) - \gamma R X]} > \frac{R}{1 - X} \quad ,$$

or

$$(\gamma\ddot{\gamma} - \gamma^2)(1 - X)^2 - \gamma^2 R^2 X \quad < \quad \gamma[\dot{\gamma}(1-X) - \gamma R X]\, R \quad ,$$

since the quantity in brackets is negative. Continuing, we find

$$(\gamma\ddot{\gamma} - \gamma^2)(1 - X) \quad < \quad \gamma\ddot{\gamma} R$$

$$\ddot{\gamma} - \frac{\dot{\gamma}^2}{\gamma} \quad < \quad \frac{\dot{\gamma} R}{1 - X}$$

$$\frac{\ddot{\gamma}}{\dot{\gamma}} - \frac{\dot{\gamma}}{\gamma} \quad > \quad \frac{R}{1 - X} \quad > \quad 0 \quad .$$

It is not clear why this condition should hold in general. If $\ddot{\gamma} > 0$, the left side may even be negative. We must conclude that weak hypotheses concerning market bias are not sufficient to explain Ben-Porath's findings. As a matter of fact, the present inequality becomes increasingly difficult to satisfy (ceteris paribus) with advancing age, since $X \equiv e^{(r+d)(t-T)}$ rises. Yet, it is only in the upper age range that the market-bias explanation is needed.

NOTES

CHAPTER I

[1]Origins of this doctrine may be traced back as far as Adam Smith, An Inquiry into the Nature and Causes of the Wealth of Nations, edited by Edwin Cannan (New York: Modern Library, 1937), p. 101, and beyond Smith, to Sir William Petty in the late seventeenth century. See Bernard F. Kiker, "The Historical Roots of the Concept of Human Capital," Journal of Political Economy, LXXIV (October, 1966), 481-499. Its modern flowering must be credited to Theodore Schultz and Gary Becker. The seminal articles were: Theodore W. Schultz, "Capital Formation by Education," Journal of Political Economy, LXVIII (December, 1960), 571-583, and "Investment in Human Capital," American Economic Review, LI (March, 1961), 1-17; Gary S. Becker, "Investment in Human Capital: A Theoretical Analysis," Journal of Political Economy, LXX (October, Supplement, 1962), 9-49.

[2]The derivation which follows is the work of Mincer, "The Distribution of Labor Incomes: A Survey." This version of the model differs from the one employed by Becker mainly in its use of continuous rather than discrete time. Cf. Gary S. Becker, Human Capital (New York: National Bureau of Economic Research, 1964), Chapter III.

[3]For the moment we may thus regard earnings and wage rates as interchangeable.

[4]According to Mincer, the latter condition is satisfied approximately in the case of American males. See Schooling, Experience and Earnings, p. 8, n. 2.

[5]Becker's early estimates imply that if college students earn approximately one-quarter as much as non-students, the income received will in fact just balance direct costs. See Human Capital, pp. 74-75. Dodge found that, on average, the part-time earnings of Canadian students greatly exceeded direct costs (Returns to Investment in University Training, Table 5.1 and 5.2, pp. 77-78). Since students sacrifice leisure as well as earnings to attend school, valuing their opportunity cost presents further problems. See Donald O. Parsons, "The Cost of School Time, Foregone Earnings, and Human Capital Formation," Journal of Political Economy, LXXXII (march/April, 1974), 251-266.

[6]Schooling, Experience, and Earnings, p. 10.

[7]We consider here only the first moments of any probability distributions connected with $W_i^*(s)$. We thus ignore the question of risk. On this point see John C. Hause, "The Risk Element in Occupational and Educational Choices: Comment," Journal of Political Economy, LXXXII (July/ August, 1974), 803-805.

[8]The required initial condition is $W_i^*(0) = W_{i0}$.

[9]Otherwise, the second-order condition $d^2 \ln W_i^*(s)/ds^2 < 0$ will not be fulfilled.

[10]See his "Underinvestment in College Education," American Economic Review, L (May, 1960), 347, or Human Capital, Chapter V.

[11]Some initial steps have been taken by Richard B. Freeman, The Market for College-Trained Manpower(Cambridge, Massachusetts: Harvard University Press, 1971), Chapter I and Chapter II. In addition to enrollment, of course, one must take into account such things as labour-force participation, deaths, retirements, and net migration.

[12]A model which incorporates demand has been tested by John F. O'Connell, "The Labor Market for Engineers: An Alternative Methodology," Journal of Human Resources, VII (Winter, 1972), 71-86.

[13]Theodore W. Schultz, "The Reckoning of Education as Human Capital," in Education, Income, and Human Capital, edited by W. Lee Hansen (New York: National Bureau of Economic Research, 1970), p. 301.

[14]A classic example is Clark Kerr, "The Balkanization of Labor Markets" in E. Wight Bakke et al., Labor Mobility and Economic Opportunity (New York: Technology Press and John Wiley and Sons, Inc., 1954); but see as well Lloyd G. Reynolds, The Structure of Labor Markets (New York: Harper and Brothers, Inc., 1951). A more recent work in this tradition is Albert Rees and George P. Shultz, Workers and Wages in an Urban Labor Market (Chicago: University of Chicago Press, 1970).

[15]Hendrik S Houthakker, "Education and Income," Review of Economics and Statistics, XLI (February, 1959), 14-17. W. Lee Hansen, "Total and Private Rates of Return to Investment in Schooling," Journal of Political Economy LXXI (April, 1963), 128-141. Giora Hanoch, "An Economic Analysis of Earnings and Schooling," Journal of Human Resources, II (Summer, 1967), 310-329.

[16]"Present Values of Lifetime Earnings for Different Occupations."

[17]A useful survey is George Psacharopoulos and Keith Hinchliffe, Returns to Education: An International Comparison (Amsterdam: Elsevier Scientific Publishing Company, and San Francisco: Jossey-Bass Inc., 1973).

[18]Schooling, Experience, and Earnings, p. 137.

[19]Barry Bluestone, Willis Murphy, and Mary Stevensen, Low Wages and the Working Poor (Ann Arbor: Institute of Labor and Industrial Relations, University of Michigan, 1973), pp. 19 ff.

[20]See Gary S. Becker, Human Capital and the Personal Distribution of Income (An Arbor: University of Michigan, 1967.

[21]According to Becker, yields decline for a number of reasons: (1) the continuing addition of a variable factor, schooling, to a fixed factor, mental and physical ability, leads to diminishing returns; (2) foregone earnings rise (faster than productivity in learning) as education accumulates; (3) the amortization period shortens; (4) the marginal utility of additional earnings falls; (5) risk aversion may rise as human capital increases. These last two arguments seem rather out of place in an income maximizing framework. As for the interest cost, it rises because of segmentation in the loans market and the need for students to resort to increasingly expensive source.

[22]See Dale W. Jorgenson, "The Theory of Investment Behavior," in Determinants of Investment Behavior edited by Robert Ferber (New York: Columbia University Press for the National Bureau of Economic Research, 1967). Note that separable utility, defined over leisure and consumption, is not sufficient to make earnings and utility maximization coincide.

[23]For the moment, however, note T.D. Wallace and L.A. Ihnen, "Full-Time Schooling in Life-Cycle Models of Human Capital Accumulation," Journal of Political Economy, LXXXIII (February, 1975), 137-156. These authors explore the extreme imperfection of no borrowing for investment purposes.

[24]Mincer adopts this orthodox interpretation, though he does briefly acknowledge the possible impact of labour-market factors. See Schooling, Experience and Earnings, p. 138.

[25]Walter Haessel and P.J. Kuch, "An Analysis of the Determinants of the Size Distribution of Earnings in Canada," University of Western Ontario, unpublished, 1976.

[26]We shall consider here only an exact specification of the model, with schooling the only form of human capital. The problems encountered when a stochastic term is present will be discussed, along with other questions of implementation, in Chapter III.

[27]See Gary S. Becker and Barry R. Chiswick, "Education and the distribution of Earnings," American Economic Review, LVI (May, 1966), 358-369; and Jacob Mincer, "Time-Series Changes in Personal Income Inequality in the United States from 1939, with Projections to 1989," Journal of Political Economy, LXXX (May-June, 1972, Supplement), S34-S66; Chiswick, Income Inequality: Regional Analysis within a Human Capital Framework (New York: Columbia University Press for the National Bureau of Economic Research, 1974); Mincer, Schooling, Experience, and Earnings, Chapter II and Chapter VI.

[28]In the present context W(0) way be interpreted as representing the individual's initial endowment of ability and human capital. Whether it is in fact uncorrelated with r and s is therefore somewhat dubious.

[29]To be precise,

$$R(r,s) \equiv 2\bar{s}E[r-\bar{r})^2(s-\bar{s})] + 2\bar{r}E[(r-\bar{r})(s-\bar{s})^2]$$
$$+ E[r-\bar{r})^2(s-\bar{s})^2] - [Cov(r,s)]^2 \quad ,$$

where E is the expectations operator. The theorem is due to Leo A. Goodman, "On the Exact Variance of Products," Journal of the American Statistical Association, LV (December, 1960), 708-713.

[30]See the well-known paper by A.B. Atkinson, "On the Measurement of Inequality," Journal of Economic Theory, VI (September, 1970), 244-263, and R. Love and M.C. Wolfson, Income Inequality: Statistical Methodology and Canadian Illustrations (Ottawa: Statistics Canada, 1976), Catalogue 13-559. A defect of the variance-of-logarithms measure is that it does not necessarily satisfy "Dalton's condition," which states that any transfer from a rich to a poor individual must register as a decline is inequality, provided the amount of the transfer is not so large as to reverse the parties' ranking in the income distribution.

It should also be recognized that the present discussion refers only to contemporaneous cross-sectional and not to lifetime inequality. Within the restricted framework of the schooling model where age-earning profiles (after graduation) are horizontal, this distinction is unimportant; but such is not always the case. See Harry G. Johnson, "Some Micro-Economic Reflections on Income and Wealth Inequalities," Annals of the American Academy of Political and Social Science, CDIX (September, 1973), 54-59, or Morton Paglin, "The Measurement and Trend of Inequality: A Basic Revision," American Economic Review, LXV (September, 1975), 598-609.

[31]Schooling, Experience, and Earnings, p. 27.

[32]Both, op. cit.

[33]Schooling, Experience and Earnings, p. 53, Table 3.3 and p. 92, Table 5.1. This material is reproduced for convenience in Appendix IIA.

[34]Schooling, Experience, and Earnings, pp. 54-55.

[35]Alan Marin and George Psacharopoulos, "Schooling and Income Distribution," Review of Economics and Statistics, LVIII (August, 1976), 332-338.

[36]We shall ignore the distribution of W_0.

[37]Nicholas Oulton, "The Distribution of Education and the Distribution of Income," Economica, XLI (November, 1974), 387-402.

[38]Ibid., pp. 388-389.

[39]The simpler and more common assumption has been that workers in different educational categories are perfectly interchangeable, according to the number of "efficiency units" they supply, along linear production isoquants. See, for example, Zvi Griliches, "Notes on the Role of Education in Production Functions and Growth Accounting," in Education, Income, and Human Capital, edited by W. Lee Hansen (New York: National Bureau of Economic Research, 1970). The more general form of (8) still rules out complementerities. Cf. Samuel Bowles, "Aggregation of Labor Inputs in the Economics of Growth and Planning Experiments with a Two-Level C.E.S. Function," Journal of Political Economy, LXXVIII (January-February, 1970), 68-81.

[40]Oulton, op. cit., p. 394.

[41]"Education and the Distribution of Earnings."

[42]One must do so, at least, within the context of the present model, which abstracts from all variation in time worked. Depreciation of human capital receives particularly detailed treatment in Sherwin Rosen, "Measuring the Obsolescence of Knowledge," in Education, Income, and Human Behavior, edited by F. Thomas Juster (New York: McGraw-Hill Book Co., 1975).

[43]Schooling, Experience, and Earnings, p. 49.

[44]Formally, observe that

$$\ln W_s = \ln Es + \ln(1-k_0);$$

$$\ln W^a_{\tilde{p}} = \ln E_s \;;$$

$$\ln W_{\tilde{p}} = \ln E_s + r_x K_p{'} + \ln(1-k_{\tilde{p}}),\ \tilde{p} < p' \leqq T \quad ,$$

where

$$K_{p'} = \sum_{t=0}^{\tilde{p}-1} k_t \quad .$$

Therefore,

$$\begin{aligned} \mathrm{Var}(\ln W_s) &= \mathrm{Var}(\ln E_s) + \mathrm{Var}[\ln(1-k_0)] \\ &\quad + 2\,\mathrm{Cov}[\ln E_s, \ln(1-k_0)] \;; \\ \mathrm{Var}(\ln W_{\tilde{p}}) &= \mathrm{Var}(\ln E_s) \;; \\ \mathrm{Var}(\ln W_{\tilde{p}'} &= \mathrm{Var}(\ln E_s) + r_x^2 \cdot \mathrm{Var}(K_{p'}) + 2r_x \cdot \mathrm{Cov}(\ln E_s, K_{p'}) \\ &\quad + 2\mathrm{Cov}[\ln E_s, \ln(1-k_{p'})] + 2r_x[K_{\tilde{p}}, \ln(1-k_{p'})]. \end{aligned}$$

If the covariances are small, $\mathrm{Var}(\ln W_{\tilde{p}})$ will constitute the minimum. Cf. Schooling, Experience, and Earnings, p. 102.

[45]Ibid., p. 103.

[46]Burton G. Malkiel and Judith A. Malkiel, "Male-Female Pay Differentials in Professional Employment," American Economic Review, LXIII (September, 1973), 693-705.

[47]Glen G. Cain, Richard B. Freeman, and W. Lee Hansen, Labor Market Analysis of Engineers and Technical Workers (Baltimore: Johns Hopkins Press, 1973); Anders Klevmarken and John M. Quigley, "Age, Experience, Earnings, and Investments in Human Capital," Journal of Political Economy, LXXXIV (February, 1976), 47-72.

[48]Edward Lazear, "Age, Experience and Wage Growth," American Economic Review, LXV (September, 1976), 548-559.

[49]Presumably, such workers do not receive any on-the-job training. See George Psacharopoulos, "Schooling, Experience, and Earnings: The Case of an LDC," Journal of Development Economics, IV (March, 1977), 39-48.

[50]Mark Blaug, "The Empirical Status of Human Capital Theory: A Slightly Jaundiced Survey," Journal of Economic Literature, XIV (September, 1976), 837.

[51]Ibid., pp. 45-47.

[52]Human Capital, pp. 11-18.

[53]Sherwin Rosen, "Learning and Experience in the Labor Market," Journal of Human Resources, II (Summer, 1972), 326-345.

[54]Ibid., p. 328.

[55]Blaug, "Human Capital Theory," p. 837.

[56]Human Capital, pp. 18-29.

[57]David Donaldson and B. Curtis Eaton, "Firm-Specific Human Capital: A Shared Investment or Optimal Entrapment?" Canadian Journal of Economics, IX (August, 1976), 462-472.

[58]Melvin W. Reder, "Gary Becker's Human Capital: A Review Article," Journal of Human Resources, II (Winter, 1967), 100.

[59]Yoram Ben-Porath, "The Production of Human Capital and the Life Cycle of Earnings," Journal of Political Economy, LXXV (August, 1967), 352-365.

[60]William J. Haley, "Human Capital: The Choice Between Investment and Income," American Economic Review, LXIII (December, 1973), 929-944. Note also the following: Eytan Sheshinski, "On the Individual's Lifetime Allocation Between Education and Work," Metroeconomica, XX (January, 1966), 42-29; Y. Comay, A. Melnik, and M.A. Pollaschek, "The Option Value of Education and the Optimal Path of Investment," International Economic Review, XIV (June, 1973), 421-435.

[61]The marginal cost function is given by

$$\begin{aligned} MC &= w/(\partial Q_H/\partial I) \\ &= (w/\alpha\gamma)I^{1-\gamma} \\ &= (w/\gamma)\,\alpha^{-1/\gamma}\,Q_H^{(1-\gamma)/\gamma}\,. \end{aligned}$$

The condition $0 < \gamma < 1$ ensures that $\partial MC/\partial Q_H > 0$. If $\gamma > 1$, $\lambda(t)$ will intersect the marginal cost function from below, and the second-order condition for a maximum will not hold. In this situation, the individual would never wish to devote any time to market work.

[62]This expression is easily derived by setting $MC(t) = \lambda(t)$, and differentiating in logs.

[63]Haley, "Human Capital: The Choice Between Investment and Income," p. 937.

[64]Ibid., pp. 937-938.

[65]Yoram Ben-Porath, "The Production of Human Capital and Time," in Education, Income, and Human Capital, edited by W. Lee Hansen (New York: National Bureau of Economic Research, 1970).

[66]Ibid., p. 139.

[67]Mincer especially has emphasized this problem. See his "Comment," in Education, Income, and Human Capital, edited by W. Lee Hansen (New York: National Bureau of Economic Research, 1970).

[68]A rigourous proof may be found in Appendix I. Ben-Poroth argues for the likelihood of increasing market bias in stating:

> The market does not make it possible to get something for nothing, so that neutral improvement in human capacity costs more than specialized improvement. . . . When there is still a large investment program ahead, it is advisable

> to emphasize devices that . . . make the individual a more efficient producer of human capital. Later, . . . the fraction of investment outlays devoted to skills that are for purposes of further investment will be smaller. ["The Production of Human Capital and Time," p. 143].

Ben-Porath thus reverts to the notion of heterogeneous human capital. Such an idea seems notably out-of-joint with orthodox human capital theory, which emphasizes the homogeneous value of self-investment.

[69]See Thomas Johnson and Frederick J. Hebein, "Investment in Human Capital and Growth in Personal Income, 1956-1966," American Economic Review, LXIV (September, 1974), 604-615.

[70]We shall of course be dealing fully with this problem in Chapters IV and V. The first to raise it seriously appears to have been Lester Thurow, "Comment," in Education, Income, and Human Capital, edited by W. Lee Hansen (New York: National Bureau of Economic Research, 1970), p. 154.

[71]Sherwin Rosen has pointed out that Mincer's investment series implicitly assume constant returns to scale. See "Income Generating Functions and Capital Accumulation," Harvard Institute for Economic Research, Discussion Paper No. 306 (unpublished), 1973.

[72]Charles Brown, "A Model of Optimal Human-Capital Accumulation and the Wages of Young High School Graduates," Journal of Political Economy, LXXXIV (April, 1976), 299-316.

[73]James J. Heckman, "Estimates of the Human Capital Production Function Embedded in a Life-Cycle Model of Labor Supply," in Household Production and Consumption, edited by Nestor E. Terleckyj (New York: National Bureau of Economic Research, 1975). Notwithstanding the title, labour supply does not enter the cited estimates in an essential way.

[74]William J. Haley, "Estimation of Earnings Profiles from Optimal Human Capital Accumulation," Econometrica, XLIV (November, 1976), 1223-1238.

CHAPTER II

PROBLEMS OF IMPLEMENTATION

Studies which seek to apply the preceding models in some way to available earnings data now make up a vast body of research. Even by 1964, efforts to compute the rates of return to various forms of education had proliferated to such an extent that Becker found it necessary to caution against "excesses" in the use of the human-capital concept.[1] The outpouring of work has continued, though undoubtedly with some important refinements.

For present purposes, there is little value in attempting to survey the quantitative results of this immense literature. Specific attention will be given to the few significant pieces of Canadian research that have appeared, and to the findings of Mincer, whose work provides a basis of comparison for the empirical results reported later in this study. Mainly, however, this chapter will examine the assorted problems of estimation and interpretation that arise in implementing the models just surveyed. Such problems must be faced, even if one holds the underlying analysis to be beyond falsification and therefore deficient as a scientific theory of individual behaviour. In the absence of further qualification, the human-capital paradigm may prove misleading even in its other, more mundane role as a framework for ex post measurement and description.

As in the preceding chapter, we shall look first at the schooling model and then at the analysis of on-the-job training. We shall consider implementation of the "general model" very briefly, since the data and methods used are of minor relevance to the current study.

THE SCHOOLING MODEL

Implementation of the schooling model appears straightforward. One has merely to add a conventional disturbance term u_i to Equation (3), so that with $W_i(0) = W_0$ for all i,

$$\ln W_i = \ln W_0 + r^e s_i + u_i \quad . \qquad \ldots .(29)$$

Regressing ln W on s over any desired cross-section of individuals then provides an estimate of r^e, the rate of return to schooling. Equation (29) assumes that r^e is the same for all members of the chosen population. In a trivial sense, therefore, the simple regression estimate portrays the mean. Equation (29) does permit individual variation in $\ln W_0$ through the additive disturbance u; but the latter, in adsorbing such variation, must remain uncorrelated with s. We shall explore in the next subsection the consequences of violating the two preceding conditions.

When Mincer applies Equation (29) to census microdata on American males, the model explains only 7% of the variance in the logarithm of annual (1959) earnings.[2] The apparent rate of return to schooling is also 7%. This value of r^e is well below the estimates of earlier American studies, which compute rates of return directly by comparing average or fitted age-earnings profiles.[3] Direct estimates for the United States typically fall in the 10-16% range.[4] Podoluk's results for Canada indicate returns of 16.3% to a high school diploma and 19.7% to a university degree.[5] In the face of such evidence, the low figure yielded by the simple-regression approach casts immediate doubt upon the validity of the schooling model.

The unimpressive value of R^2 registered by (29) is not in itself very disturbing. No one could reasonably expect the schooling model to furnish a complete description of the earnings generation process: variables other than schooling are obviously important. The simple model may nevertheless *contribute* to an adequately formulated earnings function. We must therefore look closely at the problems surrounding its implementation.

The suspected bias in the simple-regression estimate of r^e may stem from a number of econometric difficulties. These may be grouped under the following five headings: (1) individual variation in the rate of return, (2) endogeneity of schooling, (3) expectations and economic growth, (4) omission of ability and family background, (5) omission of other variables. We shall now examine each set of problems in detail.

Individual Variation in the Rate of Return

The assumption that r^e is the same for all individuals certainly places a very strong a priori restriction upon the schooling model. More generally, one might argue that individual rates of return contain a personal component v_i. Hence, we may write $r_i^e = \bar{r}^e + v_i$, as in Chapter I. For completeness, one might also recognize a personal factor w_i', governing initial earning capacity. In this case, let us write $W_{i0} = \bar{W}_0 w_i'$, so that $\ln W_{i0} = \ln \bar{W}_0 + w_i$, where $w_i \equiv \ln w_i'$. Modifying (29) appropriately, we obtain

$$\begin{aligned} \ln W_i &= (\ln \bar{W}_0 + w_i) + (\bar{r}^e + v_i)\, s_i + u_i \\ &= \ln \bar{W}_0 + \bar{r}^e s_i + u_i + w_i + v_i s_i \quad . \qquad \ldots .(30) \end{aligned}$$

Now, in the simple regression of $\ln W_s$ on s, the expected value of the estimated slope coefficient $\hat{\bar{r}}^e$ is given by

$$\begin{aligned} E(\hat{\bar{r}}^e) &= E\,[\,\sum_i s_i \ln W_i \;/\; \sum_i s_i^2\,] \\ &= E\,[\,\sum_i s_i(\bar{r}^e s_i + u_i + w_i + v_i s_i)/\sum_i s_i^2\,] \\ &= \bar{r}^e + E\,(\sum_i u_i s_i/\sum_i s_i^2) + E(\sum_i w_i s_i/\sum_i s_i^2) \\ &\quad + E\,(\sum_i v_i s_i^2/\sum_i s_i^2) \quad , \end{aligned}$$

assuming, just for the moment, that both $\ln W$ and s have been scaled in deviations from their respective means. Note that although s_i is a fixed number for any given i, it is nevertheless stochastic in the sense that the identity of the i^{th} individual will vary randomly in repeated samplings. If the simple-regression estimate is to be unbiased, the terms involving u_i, v_i, and w_i must vanish. In other words, u and w must be uncorrelated with s, and v must be uncorrelated with s^2.

The requirement pertaining to u is, of course, a standard assumption of the linear regression model. The same requirement extends naturally to w, which contributes in parallel fashion to the observable error $(u_i + w_i + v_i s_i)$. Here, we isolate w to expose analytically whatever bias may result from this one error component. In fact, some degree of bias appears highly probably, since it is difficult to believe that s and w could be independent. Factors which promote initial earning capacity seem certain to affect schooling as well. In particular, s and w may be related empirically through a mutual dependence upon ability and family background. If the relationship is positive, $\hat{r}^e$ will have an upward bias. Surprisingly, however, some theoretical arguments suggest a negative relationship. Since these arguments hinge on the precise treatment of ability and family background, they are best reserved for the subsection devoted to this topic.

Our immediate concern is the requirement that v be indepedent of s^2. Although the human-capital literature does not investigate this rather special hypothesis, it does supply abundant evidence of a general association between schooling and the rate of return. The American studies already cited document a fall in r^e, and therefore in v, as s rises. If we may thus infer a negative correlation between v and s^2, it would appear that the simple-regression estimate of $\bar{r}^e$ will contain a downward bias. This factor may help to explain the low rate-of-return estimates typically derived using the simple-regression approach.

In Canada, however, there is some evidence that rates of return increase with the level of schooling. As we have seen, Podoluk encountered higher returns among university than among secondary-school graduates. Calculations performed by Dodge for several highly trained occupations show increasing returns in three out of four cases.[6] One must therefore be alert to the possibility of an upward bias in regression estimates computed from Canadian data. The empirical work reported in Chapter III addresses this problem.

Mincer approaches the question of individual variation in the rate of return by expanding the regression model to include s^2. The derivative[7]

$$d \cdot \ln W/ds \;=\; r_0^e + 2r_1^e s$$

then provides an estimate of the marginal return to schooling. This will be declining if $r_1^e < 0$ and in Mincer's initial trials, r_1^e is indeed both negative and significant.[8] However, the significance disappears when Mincer standardizes for the number of weeks worked during the sample year.[9] On the strength of this empirical result, he concludes that rates of return computed on the basis of weekly wages are nearly constant, and that the apparent association between s and r^e is due mainly to the employment effects of schooling.[10] By implication, therefore, estimates obtained using weekly wages will be unbiased.

Yet, a problem of interpretation now arises. The rate of return, as it is normally understood, includes all the benefits attributable to schooling. Relative immunity to unemployment is possibly one of these. If so, holding weeks of work constant violates the standard concept. This procedure may well furnish an unbiased estimate, but not of the parameter we originally set out to measure. What we obtain instead--the weeks-constant rate of return--is a limited notion, with limited usefulness, perhaps, in assessing individual investment behaviour.

Blaug implicitly adopts the broad rate-of-return concept when he argues that

> Mincer's result is actually rather paradoxical. It is a fact that average weeks worked per year increase with the level of schooling. Hence, if we standardize for the numbers of weeks worked per year by calculating rates of return to schooling from weekly rather than annual earnings, the decline in rates of return to successively higher levels of schooling should increase, not decrease, the more so as there is some evidence that weekly earnings tend to be positively correlated with weeks worked per year.[11]

The paradox noted here is really a matter of confusion over Mincer's failure to distinguish between the weeks-constant and the weeks-variable rate of return. For Blaug and others, "rate of return" *means* only the latter. Empirically, the two competing measures lie rather far apart. In a pair of comparable regressions reported by Mincer, the first stands at 12%; the second, evaluated at the mean year of schooling, equals 18%.[12] Hence, one cannot justify the first measure as an approximation for the second. Whether one may legitimately hold constant weeks worked per year, or any other variable linked to schooling, is in fact a recurring problem in rate-of-return estimation. We shall meet this dilemma again later.

Right now observe that when Mincer adds s^2 to the simple-regression model, he is implicitly letting $v_i = r_1^e s_i + \tilde{v}_i$, where $\tilde{v}_i$ represents another disturbance. Substituting this hypothesis into Equation (30) yields

$$\ln W_i = \ln \overline{W}_0 + r_0^e s_i + r_1^e s_i^2 + u_i + w_i + \tilde{v}_i s_i \quad , \quad \ldots . (32)$$

with r_0^e replacing $\bar{r}^e$.[13] Estimates of r_0^e and r_1^e will now be unbiased (subject to the previous restrictions on u and w) as long as $\tilde{v}$ is independent of s^2 and s^3. If the expression for v_i succeeds in capturing the true relationship between schooling and the rate of return, there is no further reason to suspect that $\tilde{v}$ might be correlated with s, raised to any particular power. One may as well assume unbiasedness. However, because s appears in the composite error terms of (30) and (32), both models will presumably suffer from heteroskedasticity. Estimates of $\bar{r}^e$, or of r_0^e and r_1^e, will not be efficient, and the standard errors will be biased downward. This problem will not yield, moreover, to any simple transformation, since the composite disturbances are nonhomogeneous in s.

Of course, one might postulate functional relationships between s and r^e that are more complicated than the linear hypothesis examined here. An endless number of ad hoc models may be generated in this way. An alternative strategy which seems more promising is to make v a function of other variables besides schooling. One then arrives at some version of the "interactions model," described in Chapter I. In this context, the squared term appearing in (32) represents the interaction of schooling with itself. From an econometric point of view, one's goal in specifying further interactions is to explain v in such a way that the ultimate residual, $\tilde{v}$, emerges as a "clean" stochastic term, uncorrelated with any of the independent variables. Bias is thus eliminated, although the problem of heteroskedasticity remains.[14]

It is important to note, in concluding this subsection, that the issue of individual variation in the rate of return is a crucial one for human-capital theorists. Econometric difficulties aside, if the rate of return (like the velocity of money or the marginal propensity to consume) is not a stable constant when viewed in the relevant dimension--across otherwise dissimilar groups of individuals--then, the power of human-capital theory is greatly attenuated. This power lies in the notion that individual differences may be reduced to a single variable, the stock of "human capital." Multiplying the value of the stock by a simple parameter, the "rate of return," yields individual earnings. However, when the stock of human capital and the rate of return *both* depend on (possibly nondisjoint sets of) individual attributes, much of the initial clarity, even as a descriptive framework, is lost. The interactions model, even though it follows quite naturally from Becker's supply-and-demand framework, violates the spirit of orthodox human-capital analysis.

Endogeneity of Schooling

As soon as one pays explicit heed to the market processes which underlie the statistical relationship between schooling and earnings, it becomes apparent that schooling need not be an exogenous variable. On the demand side of the labour market, schooling determines earnings; but on the supply side, where individuals make investment decisions, earnings determine schooling.[15] Equation (29) may thus contain a degree of simultaneity bias.

Formally, we may think of the following static equilibrium system:

$$L_{(s)}^{dem} = L^{dem}(\bar{W}_0, \bar{W}_1 \cdots, \bar{W}_s, \cdots, \bar{W}_n; s, z_1) \quad \ldots . (33)$$

$$L_{(s)}^{sup} = L^{sup}(\bar{W}_0^*, \bar{W}_1^*, \cdots, \bar{W}_s^*, \cdots, \bar{W}_n^*; s, z_2) \quad \ldots . (34)$$

$$\bar{W}_s^* = f(\bar{W}_s) \qquad s = 0, 1, \cdots, n \qquad \ldots.(35)$$

$$L^{dem} = L^{sup} \quad . \qquad \ldots.(36)$$

The first two equations are a demand and a supply function respectively. As in the preceding text, L's stand for aggregate numbers of individuals, bars over the W's indicate means, and asterisks denote ex ante variables.[16] Two stochastic elements, z_1 and z_2, allow for maximizing errors and other, unspecified influences. The third equation links observed and expected wages.[17] The last is an equilibrium condition. Substituting into it from (33), (34), and (35), we obtain the locus[18]

$$M(\bar{W}, s, z_M) = 0 \quad , \qquad \ldots.(37)$$

where z_M is a function of z_1 and z_2.

Now, the schooling model imposes upon this locus of equilibrium points a particular functional form--that displayed in Equation (29). Using microdata instead of grouped observations, we must of course insert the individual disturbance variable u in place of z_M. However, nothing in the derivation of the schooling model requires that we solve (37) for W. We could as well have written

$$s_i = \frac{1}{r^e} \left\{ - \ln W_0 + \ln W_i - u_i \right\} , \qquad \ldots.(38)$$

which also yields an estimate of r^e. In general, this estimate will not agree with one obtained from Equation (29).[19] Since (38) and (29) both implement the fundamental postulate of equal present values, it is not clear a priori which one the researcher should employ.

This simple view of the endogeneity problem is reinforced when we consider explicitly the individual's optimizing behaviour. Recall that in Chapter I we derived the optimality condition $dW_i^*/ds = W_i^* r_i$. By the chain rule,

$$dW_i^*/ds = (dW_i^*/dW_i)(dW_i/ds) \quad .$$

Let us suppose that $W_i = W_0 e^{r^e s_i} + u_i$, where r^e is the "true" rate of return available in the market. Then,

$$dW_i/ds = r^e W_0 e^{r^e s_i} + u_i \quad .$$

We noted in Chapter I that the second-order condition for optimality will be satisfied only if $d^2W_i^*/ds^2 < 0$. Assuming that $dW_i^*/dW_i = f'(W_i) > 0$, we can meet this requirement by making r^e a declining function of s. Let us do so implicitly in order to keep the ensuing algebra relatively simple.[20]

The preceding results, together with Equation (35),[21] now imply that for optimality to hold

$$f'(W_i) \cdot (r^e W_0 e^{r^e s_i} + u_i) = f(W_i)\, r_i \quad ,$$

or

$$\ln[f'(W_i)] + \ln r^e + \ln W_0 + r^e s_i + u_i = \ln[f(W_i)] + \ln r_i \quad .$$

$$\therefore s_i = \frac{1}{r^e}\left\{- \ln W_0 + \ln[f(W_i)] - u_i + \ln[r_i/r^e] - \ln[f'(W_i)]\right\} \quad .$$

If expectations coincide with existing market opportunities, $f(W_i) = W_i$, $f'(W_i) = 1$, and for the marginal investor at least, $r_i = r_i^* = r^e$. In this case, (39) reduces to (38).

If this analysis is correct, (29) and (39) form a simultaneous system in which s depends *negatively* upon u.[22] Single-equation estimates of (29) may, therefore, yield values of $\hat{r}^e$ that are biased downward. Results reported by Griliches suggest that the downward bias may be as much as 40%.[23] If so, we cannot dismiss the problem lightly.

Defenders of the single-equation approach may nevertheless argue that in cross-sectional data schooling is a predetermined variable. Current levels of schooling are the product of decisions taken in the past on the basis of expectations formed in the past. These expectations may depend, in turn, upon market conditions prevailing in periods even further removed from the present. In the case of some older workers, we may thus be dealing with time spans as long as 40 or 50 years. Under such circumstances, a direct *behavioural* link between schooling and current wage rates is clearly impossible.

We know, however, that wage structures evolve rather slowly. At the same time, individuals may not be totally unsuccessful in forecasting the future. We may, therefore, encounter a significant *statistical* relationship between schooling and current wages. As Griliches explains, "To the extent that the 'errors' (from the point of view of us as observers) in the ex-post and ex-ante earnings functions are correlated, they will be 'transmitted' to the schooling equation and induce an additional correlation between schooling and these disturbances."[24] The result will be simultaneity bias. In the formal model sketched here, the required "transmission" role is performed by (35). That this equation may depict correlation rather than causality is of no great importance.

It might further be argued that schooling is not dependent upon earnings because it is not, to any significant degree, the subject of optimizing behaviour. According to this view, such things as tastes, socioeconomic background, and the decisions of parents serve as the main determinants of individual schooling. Actually, parental decision-making need not affect our earlier analysis. If parents are altruistic and as well informed as their children, they may plan to maximize children's lifetime earnings in just the way we have previously hypothesized.[25] It may be that a great many factors--tastes and socioeconomic background among them--determine schooling; but if the set of determinants excludes earnings, a dilemma appears. Without some link between schooling and earnings, there is no mechanism for disequilibrium adjustment.

If levels of schooling observed in cross section *are* predetermined, the supply functions of the preceding market model describe vertical lines. With demand functions given, the resulting locus of short-run equilibrium points may look like (29), or it may not. At best, we have a problem of interpretation. The rates of return derived using (29) are themselves short run in character. More precisely, they are the rates a current investor in schooling might earn if the current wage structure were to persist. They are not necessarily the long-run rates of return envisioned in deriving the ex ante version of the schooling model.

The nature of the dilemma should now be fully apparent. If we wish to interpret our regression coefficients as long-run, equilibrium rates of return, we must recognize the endogeneity of schooling; but if we recognize the endogeneity of schooling, we must concede that our regression coefficients may harbour simultaneity bias. In upholding the schooling model as a behavioural theory, we encounter an econometric problem.

The obvious solution is to adopt a simultaneous-equation approach. Whatever method one chooses, its success will ultimately depend on finding exogenous variables which perform well as predictors of individual schooling. Census data do not seem especially rich in this regard. The present study will not explore the endogeneity question further, though it remains an important topic for future research.

Expectations and Economic Growth

One might gather correctly from the brief and somewhat tentative remarks of the preceding subsection that the human-capital literature has very little to say on how expectations are formed. Freeman, who has written most on the topic, distinguishes three general influences: current wages, their rates of change, and nonwage factors.[26] However, in his empirical investigations, he takes only current wages as his proxy for expected lifetime earnings.[27] He thus assumes what might be called "myopic" expectations. The standard rate-of-return studies[28] ignore expectations almost completely, leaning implicitly toward an ex-post interpretation of results.

From an econometric standpoint, the most important general question we have to consider is whether the practice of ignoring expectations leads to a misspecification of the earnings function through the omission of significant explanatory variables.[29] It might be argued that if "conditions" and recent economic trends--in a particular region, at a particular time--seem to favour a particular level of schooling as an investment goal, we should then observe in our cross-section data a larger number of individuals than would normally occupy the given age-schooling cohort. If, in addition, workers belonging to the various cohorts are not perfect substitutes for one another in production, we might also observe a lower than average wage for the given cohort.[30] This wage disparity may follow the group in question throughout its life history. To allow for the possibility, one might consider adding age and region of schooling to the previous earnings function. According to the argument just outlined, these variables would represent the state of expectations prevailing at the time and in the place educational decisions were made.

The trouble with the foregoing interpretation is that it seems to preclude our saying anything in general about the effects of age and region of schooling. Suppose we learn, for example, that fifty-five year old high-school graduates from British Columbia enjoy an earnings advantage over other fifty-five year old Canadians at the same level of education. If we adhere strictly to our state-of-investor-expectations hypothesis, we cannot make any predictions whatever concerning British Columbia high-school graduates who reach fifty-five years of age at some point in the future. Age and region merely flag once-and-for-all disturbances in the pattern of educational investment.

Still, if these variables, representing transitory influences, are ignored, their omission may bias any attempt to measure the "normal," "permanent," or "long-run" rate of return. According to the familiar errors-in-variables argument, the bias will be toward zero--in the present case, negative. Age and region combat it by serving as proxies for the swings in expectations which produce "errors" (from our point of view) in the schooling variable. These errors, if we may refer to them as such, arise not from statistical measurement, but from the "mistakes" individuals make because they cannot foresee market developments.

Whether or not individuals foresee and act upon detailed changes in the educational wage structure, they may still take into account general wage advances due to economic growth. This factor gives rise to another problem in estimating both the ex-ante and the ex-post rate of return to schooling. Recall that in deriving the basic schooling model, we assumed that annual wage rates would remain constant throughout the individual's working life. The more realistic assumption--that real wages will grow exogenously over time--requires some modification of the previous result.

Let us suppose that wages are expected to rise according to the growth formulae

$$W(s,\ t) \ = \ W(s,\ 0) \cdot e^{g_s^* t} \qquad s = 0,\ 1,\ \cdots,\ n \ ,$$

where W(s, t) measures the reward to s years of schooling at time t, and the g_s^* stand for expected rates of growth, allowed for the moment to differ by level of schooling. If we again enforce the equalization of discounted lifetime earnings, it is a simple matter to show that

$$W(s,\ 0) \ = \ \frac{r^e - g_s^*}{r^e - g_0^*} \cdot W(0,\ 0) \cdot e^{(r^e - g_s^*)s} \qquad \ldots .(40)$$

replaces (3) as the equilibrium condition[31] at t = 0. Equation (40) indicates how the equilibrium wage structure may become distorted when expected growth rates differ. In general, individuals trade present earnings for future gain. When expected growth rates are all equal or cannot be distinguished on account of great uncertainty, (40) reduces to

$$\ln W(s,\ 0) \ = \ \ln W(0,\ 0) + (r - g^*)\, s \qquad \ldots .(41)$$

after letting $g_0^* = g_1^* = \cdots = g_n^* = g^*$ and taking logarithms.

If we how attempt to estimate (41) using a regression equation like (30), we encounter an elementary sort of identification problem. The slope coefficient we obtain measures $(r^e - g^*)$ rather than r^e. If we recognize depreciation (in effect, negative growth), it measures $(r'^e + d - g^*)$.[32] To "identify" r'^e, we must have some independent estimate of $(d - g^*)$. Even if we are interested only in the net rate of return (r^e), forgetting about growth may lead us to underestimate the value of this parameter.

Miller appears to have been the first to call attention to the problem of underestimation.[33] He observed that economic growth causes the lifetime earnings profiles of successive age cohorts to shift upwards. At any given time, the lowest of these profiles will therefore belong to the oldest members of the population. As a result, when we draw a *cross-section* age-earnings profile, we obtain a curve that is flatter than any of the lifetime earnings trajectories we are in fact trying to represent. This flattened cross-section profile yields an underestimate of the return to schooling. In *Human Capital,* Becker recognized the problem and computed separate rates of return for each of several assumed rates of economic growth.[34] Whether one computes the rate of return directly from age-earnings profiles or adopts the regression approach favoured by Mincer, a reasonable assumption concerning g^* (or its ex-post realization g) seems the only possible recourse in most cases.[35]

The situation is different when the researcher has at his disposal a series of repeated cross sections. Then it is possible to estimate g by following the respective cohorts over some period of actual calendar time. In this manner, Johnson and Hebein arrive at exogenous growth rates in the 3-5% range.[36] Haley's estimates are a little lower, falling roughly in the 2-4% interval.[37] These figures, imprecise as they are, give some idea of the correction one must think of applying to single-cross-section estimates based on Equation (41).

Omission of Ability and Family Background

Without question, the most persistent challenge to the schooling model has come from the broad stream of empirical research which seeks to measure the effect on earnings of ability and family background. Embedded in the resulting controversy are at least three major issues. One concerns the *relative* importance of schooling, versus background and ability, in explaining the level and distribution of earnings.[38] Another concerns the problem of "screening" and the extent to which education truly enhances worker productivity.[39] The last has to do with estimating, in an unbiased manner, the *absolute* importance of schooling--that is to say, the rate of return. This final issue is the one which has provoked the greatest argument and the one which bears most heavily upon the work of the present study.

The core of the problem is simple and well known. From the very beginning of the human-capital era, it has been conceded that if background and ability exert a direct influence on the level of earnings, neglecting their contribution may lead one to overestimate the impact of education.[40] Earnings differentials due in fact to superior abilities and to the high socioeconomic standing of parents will be credited mistakenly to the additional schooling which these favourable attributes tend to encourage. In more precise terms, the omitted-variable formula of econometric analysis states (using the standard "dot" notation) that

$$\hat{\beta}_{Ws} = \hat{\beta}_{Ws \cdot a} + \hat{\beta}_{Wa \cdot s}\ \hat{\beta}_{as} , \qquad \ldots .(42)$$

Here, $\hat{\beta}_{Ws}$ corresponds to $\hat{r}^e$, and a stands for some ability or background variable excluded from the simple model. The degree of bias in the zero-order coefficient $\hat{\beta}_{Ws}$ depends on the direct influence of a on earnings ($\hat{\beta}_{Wa \cdot s}$) and on the strength of the association between a and schooling ($\hat{\beta}_{as}$). If both are positive, so is the resulting bias.

Interestingly enough, it is not clear a priori that $\hat{\beta}_{as}$ must be greater than zero. In the Ben-Porath model, background and ability may be thought to affect the parameters H_0 (initial human capital or earning capacity) and α (personal efficiency in the production of further human capital). Yet, as we noted in Chapter I, these two factors influence the period of specialization in opposite ways. If s measures, at least roughly, the period of specialization, and if a is a variable which governs both H_0 and α, then it follows that $\hat{\beta}_{as}$ may be negative. Empirically, of course, there is general agreement that s is positively associated with the standard proxies for ability and family background.[41] Given the model, one must conclude either that α (the positive influence) is more important than H_0 or that the standard proxies favour it on average. At the same time, one might ask whether financing imperfections associated with background, but ignored by the model, are not an important factor in the empirical result.

In any event, Mincer points out that if ability or background affects earnings *only* by way of additional schooling, $\hat{\beta}_{Ws}$ will suffer no bias.[42] Although $\hat{\beta}_{as}$ may be positive, $\hat{\beta}_{Wa \cdot s} = 0$. In this case, schooling is an essential input used for converting latent advantages into marketable skills. Hause, on the other hand, has argued that ability and schooling are really complements.[43] As such, they enter the earnings function interactively. Under these circumstances, not only is $\hat{\beta}_{Wa \cdot s}$ nonzero, but its value depends also on the particular level at which s is held constant.

The consensus among American studies has been that where a measures IQ or some other test score, $\hat{\beta}_{Wa \cdot s}$ is small but statistically different from zero. Though results vary, the typical estimate of bias in $\hat{\beta}_{Ws}$ is rather small as well.[44] Griliches and Mason,[45] for example, find it to be on the order of 11-15%. Dodge reaches a similar conclusion with respect to a sample group of Canadian professionals, although his results are by no means unambiguous.[46] In the extreme, Behrman, Taubman, and Wales[47] obtain a bias estimate as high as 62% using a sample of male twins.

Elsewhere, Taubman and Wales come to the rather distressing inference that the percentage bias varies across age cohorts.[48] If so, we cannot think of applying any overall "ability correction" to the zero-order coefficient $\hat{\beta}_{Ws}$. Griliches has reinforced this view with the general observation that a standard percentage adjustment must have $\hat{\beta}_{Ws} = \hat{r}^e$ as its denominator.[49] Yet, $\hat{r}^e$ is bound to vary, perhaps widely, depending on the group of individuals in question and on the precise specification of the estimating equation. There is no reason to believe that the absolute bias (the numerator) will vary in order to keep the percentage bias constant. Finally, to compound the uncertainty, Welch has argued that if s and a, our proxies for "education" and "ability," harbour a significant degree of measurement error, even the direction of bias in $\hat{\beta}_{Ws}$ is indeterminate.[50]

Because the census data employed in the present study offer no reasonable proxies for ability or socioeconomic background, we shall not inquire further into the preceding difficulties. Although the results displayed in Chapters III and V remain very useful, they cannot, on this account, fully escape qualification.

Omission of Other Variables

It was noted in Chapter I that Mincer's "reduced-form equation"--the schooling model--contains no exogenous variables from the demand side of the labour market. It is now appropriate to inquire whether the omission of such variables might not also bias the estimated return to schooling, just as in the case of ability and family background. Over the years, interindustry studies[51] have isolated a number of factors which seem to be important in determining wage levels. These include working conditions, unionization, capital intensity, concentration, profitability, the growth rate, and plant size. If the schooling of the typical worker in an industry happens to be correlated systematically with any of the preceding variables, bias should theoretically ensue.

Whether an empirical bias does in fact arise through the omission of industry variables remains to be discovered. The inter-industry studies do provide some evidence of an interaction among wages, schooling, and other variables. Weiss detects a relationship, first, between schooling and industry concentration, and second, between schooling and the level of unionization.[52] Haworth and Rasmussen find that median labour-force schooling, adjusted for quality, adds significantly to the explanatory power of their inter-industry wage regressions.[53] However, because they focus upon the coefficients of the industry variables and not upon the one associated with schooling, their results offer little help in answering the question posed here.

Most authors of the human-capital school have simply ignored the problem, but Hanoch has taken explicit pains to deny its relevance. He argues that

> . . . a high degree of mobility exists among occupations and among industries, and this mobility depends strongly on schooling and age. . . . In other words, an individual who completes more years in school would expect to move upward in the occupational scale and perhaps to work in a better-paying industry. This is in fact the main channel by which he can realize returns on his additional investment in education. . . . As a result, it was decided to exclude occupation and industry variables from the equations and thus avoid serious biases in the estimated coefficients of schooling which, after all, are the target estimates of this analysis.[54]

There are two related points to consider here. One has to do with mobility; the other, with deciding which variables are to be held constant and which are to be left free in estimating the return to schooling. Let us deal with each of these issues in turn.

Leaving aside for a moment the specific problem of occupation, one must concur that if mobility enforces long-run equilibrium (as seen by investors in human capital), then industry variables require no separate consideration. The schooling model represents the only possible wage structure, and any long-run adjustment of factor proportions needed to maintain it will arise without fail. As we observed in Chapter I, human-capital theorists rely completely on this assumption. Whether labour mobility in the real world is actually sufficient to keep the wage structure near long-run equilibrium *at whatever point one might happen to choose for cross-section study* is nevertheless an open question. "Temporary" disequilibrium present at the time a cross-section is gathered may give a false picture of the equilibrium wage structure. Sustained market imperfection may do the same. However, if industry variables capture both kinds of distortion, including them in the earnings function may eliminate these two potential sources of bias.

We now come to the second issue. It is Hanoch's contention that including industry variables (perhaps as a set of dummy regressors) will *cause* a bias in the schooling coefficient. He argues that one cannot legitimately measure the rate of return to schooling with industry of employment held constant. The two variables, industry and schooling, are related, he says, hierarchically, with the latter being the primary determinant of wages. One may infer that the use of both in the earnings function will give rise to a problem of redundancy somewhat akin to multicollinearity. The schooling coefficient, or rate of return, will be underestimated as a result.

It is noteworthy that in a similar situation involving weeks worked, Mincer chose to include the additional variable.[55] Hanoch, in comparison, allows schooling "the benefit of the doubt." He assigns to it all the earnings covariance mutually explained by schooling and industry. In the absence of a properly specified multi-equation model to predict the worker's industry of employment, there is unfortunately no clear test with which to refute this procedure. Yet, in the face of Hanoch's rather extreme assumption, it seems only prudent to investigate the alternative case. It may turn out that including industry of employment adds little to the explanatory power of the earnings function and leaves the schooling coefficient substantially unaffected. From the latter outcome, if it should transpire, one might conclude that industrial mobility is not an important factor in realizing the returns to education. We shall come back to this point in assessing the empirical results of Chapter III.

Meanwhile, let us concede that Hanoch's argument gains considerable force when applied in the case of occupation. Without question, occupation and schooling are intimately connected. Empirically, however, the strength of any statistical association will depend on how occupations are defined. A classification scheme grounded principally in education will obviously lead to a higher correlation than one based upon industrial function. Disequilibrium and "permanent" imperfection in the occupational wage structure are also possible. Thus schooling and occupation will not be completely interchangeable in accounting for the variance of earnings. As in the case of industry, it appears worthwhile to include the questionable factor, occupation, in the earnings function, at least on a provisional basis, to establish the degree of statistical overlap with schooling and to limit thereby the range of doubt concerning the independent impact of each variable.

It is, finally, somewhat surprising in view of Hanoch's treatment of industry and occupation that he does not recognize geographic mobility as a proximate source of the return to education. By computing separate rates of return for Americans in the North and South, he in effect holds place of residence constant.[56] Yet, one could presumably argue, in the manner of the previous quotation, that highly schooled individuals obtain part of the return on their investment through migration to (or residence in) high-wage areas.

Schooling and migration (residence) may be related hierarchically in the same way as schooling and industry.

On the other hand, place of residence may exert its own influence on earnings. Geographic immobility may prevent the equalization of wages in the long and in the short run. In some resource-rich areas labour may succeed in bargaining economic rents away from rival factors. Whatever the precise circumstances, it is unlikely that all of the return to living in a particular place will be attributable in the end to schooling. Part will be due to the residence decision, just as part of the return to industry and occupation will be due to investment in job search and career planning. Hanoch seems justified therefore, despite the apparent inconsistency of his approach, in holding place of residence constant. We shall likewise insert this variable, along with industry and occupation, in the expanded earnings functions of Chapter III.

In each case, the rationale for inclusion is, first of all, to capture any fundamental disequilibrium present in the earnings structure, as seen from the perspective of the schooling model. Forming part of any *apparent* disequilibrium may be the equalizing differentials thought to compensate for various nonpecuniary items in the employment setting. These differentials are the result, not of market imperfection, but of markets functioning in a smoothly competitive manner. Even so, the three variables in question may assist in measuring the pecuniary rate of return to schooling by impounding statistically the wage differentials associated with nonpecuniary factors.

Industry, occupation, and place of residence would appear to be reasonable proxies for many of the factors one could name. The use of these variables seems especially warranted in view of existing evidence which reveals a significant correlation between nonpecuniary gains or losses and schooling.[57] Bias in the schooling coefficient is otherwise a strong possibility.

Besides industry, occupation, and place of residence, there are a number of census variables one might think of adding to the earnings function on an experimental basis. The list includes: marital status, family membership, family size, rural or urban residence, period of immigration, official language, ethnic group, religion, place of highest grade in school, major source of income. In the case of each variable, it is a simple task to formulate one or more reasonable hypotheses which define some link with earnings. We shall leave details of such hypotheses to Chapter III. Here, it is sufficient to note that if any of the preceding variables are correlated with schooling, their inclusion or omission is bound to affect the schooling coefficient. Finding out how the latter responds each time a new variable is added to the earnings function would appear to be a worthwhile undertaking. The information derived from this empirical exercise should place us in an improved position to judge the compact specification favoured by most human-capital theorists.

Normally part of this specification, though an "omitted variable" from the standpoint of the schooling model, is time worked. Since Chapters IV and V deal at length with the issues surrounding time worked, we need not discuss them here, except to mention a few brief points which will shortly become significant. First of all, as soon as we consider variation in time worked, it is necessary to distinguish between the wage rate and earnings. So far we have used these concepts interchangeably. Now let us make W stand only for the periodic wage, Y for annual earnings, and h for the number of periods worked per year. If W and h are unrelated, we might specify $Y_i = W_i h_i u_i'$, or $\ln Y_i = \ln W_i + \ln h_i + u_i$, where $u_i = \ln u_i'$. According to this simple argument, the elasticity of earnings with respect to time worked should equal unity.

If we look upon the schooling model as explaining W, substitution from (29) implies

$$\ln Y_i = \ln W_0 + \bar{r}^e s_i + (1+\theta) \cdot \ln h_i + u_i \quad , \qquad \ldots . (43)$$

with $\theta = 0$. In Mincer's research, $\hat{\theta}$ is nowhere constrained and always turns out to be significantly greater than zero.[58] Hence, either the estimation procedure is biased in some way, or wage rates in fact depend upon time worked. These are the questions we shall explore in Chapters IV and V. For now, we may generally observe that if the wage rate and time worked both depend on personal attributes (other than schooling) for which time worked is an effective proxy, then it is reasonable that $\hat{\theta}$ should be nonzero. The introduction of variables more closely portraying the attributes in question should cause its value to decline. Still, under certain conditions,[59] θ may continue to exceed zero if an overtime premium figures heavily in the typical remuneration formula.

In Mincer's regression estimates, time worked is essentially an ad-hoc insertion. Appended to the human-capital earnings functions, it greatly increases their explanatory power.[60] Actually, time worked proves only a little less important than schooling in the overtaking set, adding about 0.27 to the value of R^2.[61] Wherever Mincer achieves his most impressive statistical results--in those equations for which the R^2 exceeds 0.50--he does so through the insertion of the time-worked variable. We shall test its performance, using Canadian data and the same, single-equation techniques, in Chapter III.

THE POSTSCHOOL INVESTMENT MODEL

Somewhat ironically, Mincer bases his own objection to the schooling model on an omitted-variable argument. He points out in *Schooling, Experience, and Earnings* that when individuals spend their time acquiring formal education, they ineluctably sacrifice, along with income, the opportunity to engage in alternative methods of human-capital accumulation.[62] Time devoted to schooling obviously limits the time available for such things as on-the-job training and learning by doing. Among individuals of a given age, one would consequently predict an inverse correlation between years of school attendance and the quantity of postschool investment. Therefore, in omitting postschool investment from the earnings function, we bias downward the estimated return to schooling. In this fashion, Mincer accounts for the small coefficient thrown up by the simple regression model.

Correcting its probable bias means finding a way to measure postschool investment. Though individuals may sometimes use post-school leisure to augment their human capital, we normally associate investment activity with time spent on the job. Cumulative work time or "experience" thus measures potential investment. Measuring realized investment involves two steps. The first is to to estimate years of experience; the second is to specify the lifetime investment profile. These problems occupy the next two subsections. The third and final subsection in this part surveys very briefly the results obtained by holding postschool investment constant, first in a parametric, and then in a nonparametric manner.

Estimating Years of Experience

Because ordinary census data provide no direct information on work histories, Mincer chooses as a proxy for experience the individual's current age, minus his age at school leaving.[68] The latter equals mean years of school attendance for those in the individual's schooling category, plus five years, the presumed age at school entry. In effect, Mincer assumes that, between the end of formal schooling and retirement at age sixty-five, individuals never take a holiday from the labour force or become unemployed.

In the case of prime-age males, whose commitment to the labour force is seldom interrupted, this assumption is perhaps admissible as a first approximation; but in the case of women, whose labour-force participation tends to be irregular and discontinuous, it is highly inappropriate. For this reason, Mincer excludes women from his data set.[64] The present study adopts the same expedient.

Problems in applying Mincer's proxy to a sample consisting entirely of males nevertheless remain to be overcome. Although prime-age males seldom desert the labour force, they clearly differ with respect to lifetime unemployment. Such differences are an obvious source of measurement error. Hence, if we use the suggested proxy in a linear regression and make the simplest assumption--that its errors are uncorrelated with any of the accompanying variables or stochastic terms--standard econometric reasoning asserts that the coefficient of "experience" will have a downward bias. Blinder makes the additional claim that if schooling is the only other independent variable in the regression, its coefficient will have an *upward* bias.[65] In fact, this contention is false. It is shown in Appendix IIB that as long as schooling and experience are negatively correlated, the coefficients of *both* variables will be underestimated.

Actually, as Blinder points out, the standard econometric proof does not quite fit the case under discussion. Owing to the way in which the lifetime investment profile is usually specified (see below), the experience proxy does not enter the earnings function as a single, linear regressor. Furthermore, its measurement error does not have an expectation equal to zero. Because actual experience may fall short of but never exceed experience as defined by the proxy, the embedded errors should all be nonnegative. In making this comment, however, Blinder fails to notice that the second of two terms used in computing the proxy--that is, age at school leaving--may itself be measured with error. Hence, the discrepancy between actual and imputed experience need not *always* be positive. In any event, a positive expectation does no more than alter the constant term in the regression equation.[66]

Apart from the two difficulties mentioned by Blinder, there are other considerations which may render the standard econometric proof inapplicable. One is the possibility that errors in the experience proxy may be correlated with the level of schooling. If the latter affects cumulative lifetime unemployment--the most obvious source of measurement error--in the anticipated direction (negatively, in other words), we must presume an negative correlation of some unknown magnitude. A further possibility is that errors in the proxy may be correlated with the true level of experience, being thus heteroskedastic. It is only reasonable to suppose that cumulative unemployment will increase along with experience over the individual's lifetime. This problem, however, will not upset any qualitative conclusions. A final consideration is that schooling may be measured with error. We must concede this possibility, if only because the data are often reported in class intervals rather than by specific year.

Under the foregoing circumstances, we cannot predict the direction of bias in either the schooling or the experience coefficient a priori. Empirically, Malkiel and Malkiel[67] (also cited by Blinder) find that the schooling coefficient is biased upward (as Blinder guessed) by about 12% of its "true" (estimated) value. On the basis of the argument given in Appendix IIB, one would have to infer that this upward bias is a result of the suspected inverse correlation between schooling and the error in the experience proxy. As initially forecast, the experience coefficient is biased downward--by about 19%.

Specifying the Investment Profile

After settling on a proxy for cumulative work time, Mincer proceeds to the second obstacle in estimating postschool investment--that of determining the proportion of work time devoted in each period to the acquisition of human capital. In terms of the theoretical discussion presented in Chapter I, the problem is to specify the form of $k'(p)$, where as before, p is the year of experience. Mincer advances two hypotheses:

$$k'(p) = k'_0 - k'_0 \cdot p/T' \qquad 0 \leq p \leq T' \qquad \ldots .(44a)$$

$$k'(p) = k'_0 \cdot e^{-\beta p} \qquad \ldots .(44b)$$

Here, β is a positive constant; and though T' *may* be the date of retirement, it is more generally the date at which gross investment falls to zero. The first equation is a linear relationship in which the propensity to invest falls from $k'_0 \leqq 1$ at $p = 0$ to zero at $p = T'$. The second equation is a declining exponential which originates at k'_0 but remains positive at $p = T'$.

Both specifications seem to have been chosen for their tractability in estimation, since neither of them closely resembles the theoretical investment profile yielded by the income-maximization model. Haley's version,[68] for example, implies a functional form with the general properties of a third-order polynomial in p. Whether (44a) or (44b) might succeed in approximating such an investment profile is difficult to say. Both satisfy the minimum a priori requirement that k'(p) decline over the life cycle, but in all other respects, the two equations are ad hoc. The exponential hypothesis, (44b), is further suspect insofar as it does not constrain k'(p) to zero at any point.

To derive estimating equations, one may substitute (44a) and (44b) alternately into the continuous time version of (16'), namely:

$$\ln W_p = \ln W_0 + (r'^e - d)\, s \int_0^p [r^x k'(t) - d]dt$$

$$+ \ln [1 - k'(p)] \quad . \qquad \ldots . (45)$$

Performing the integration and expanding the last term in a Taylor series up to the quadratic yields

$$\ln W_p = a + b_1 s + b_2 p - b_3 p^2 \quad , \qquad \ldots .(46a)$$

where

$$a = \ln W_0 - k_0' (1 + k_0'/2) \qquad b_2 = r^x k_0' + k_0'(1 + k_0')/T' - d$$

$$b_1 = r'^e - d \qquad b_3 = -r^x k_0'/2T' - k_0'^2/T'^2$$

in the first case, and

$$\ln W_p = a + b_1 s + b_2 e^{-p} + b_3 e^{-2p} - d \cdot p \quad , \qquad \ldots .(46b)$$

where

$$a = \ln W_p + r^x k_0'/\beta \qquad b_2 = -r^x k_0'/\beta - k_0' - d$$

$$b_1 = r'^e - d \qquad b_3 = -(k_0')^2/2 \quad ,$$

in the second. The linear hypothesis thus leads to a quadratic estimating equation, and the exponential hypothesis, to a form known as the "Gompertz curve."

The nature of the quadratic specification is best appreciated by inserting the variables which underlie the experience proxy. If we let A stand for age, then according to the definition in the last subsection, $p = A - s - 5$. In attempting to estimate (46a), we are thus dealing with

$$\ln W_p = a + b_1 s + b_2 (A - s - 5) + b_3 (A - s - 5)^2$$

$$= (a - 5b_2 + 25b_3) + (b_1 - b_2 - 10b_3)s + b_3 s^2$$

$$+ (b_2 - 10b_3)A + b_3 A^2 - 2b_3 As \quad . \quad \quad . . . . (47)$$

This result differs somewhat from the traditional earnings profile, an equation in s, A, and A^2. Mincer argues that the traditional form provides an underestimate of the return to schooling, inasmuch as $b_2 > 0$.[69] Actually, as the preceding algebra demonstrates, the relevant condition is that $b_2 + 10b_3 > 0$ --a requirement that nevertheless appears equally true in practice. Secondly, the traditional form ignores a potentially important interaction between schooling and age.

One can see from Equation (47) that Mincer's quadratic estimating function really contains two novelties: the use of the interaction term and a restriction on its coefficient. The latter is constrained to equal $-2b_3$. In view of the concealed restriction, it cannot be assumed that adding the interaction variable--through the use of p and p^2 rather than A and A^2--will improve the fit of the equation. However, if one were to estimate the second line of (47) explicitly, it would be possible to test the validity of the restriction and, indeed, the significance of the interaction term when its coefficient is unconstrained. Mincer does not examine these two minor statistical questions.

As for the problem of bias in the estimated return to schooling, it is likely true that substituting a quadratic in p for the traditional quadratic in A will increase the schooling coefficient; but this effect is purely mechanical. It *must* occur, given the way in which p has been defined. An independent measure of experience might not lead to the same result. In any event, one should be careful not to confuse the downward bias flowing from the alleged misspecification of the earnings function (the use of A instead of p) with that arising from the outright omission of experience. Since A and p are bound to be highly correlated, the second form of bias is potentially the more severe.

The exponential hypothesis, implemented through (46b), does not lend itself so easily as the quadratic specification to comparison with the traditional earnings function. We shall be content, therefore, merely to review its performance in estimation.

The Empirical Outcome

Before we examine Mincer's quantitative results, note that while (46a) yields to the standard linear-regression approach, (46b) is more demanding. Because of the very large sample Mincer employs, highly sophistocated nonlinear techniques are no doubt impractical. Understandably, he resorts to direct trial and error. Assigning different values to β, he computes a series of linear regressions and chooses the one (or the pair, as it turns out) with "the highest R^2 and the most plausible coefficient [s]."[70]

Unfortunately, we cannot look at Mincer's reported regressions (see Appendix IIA) and compare precisely the empirical performance of his two competing hypotheses. No two equations differ only in this one aspect. It does appear that the exponential form holds an advantage, although the difference--perhaps 2 or 3 percentage points in the value of R^2--is rather slight.[71] Both models explain roughly 30% of the variance in annual earnings.

The real advantage of the exponential form lies in its ability to identify the parameters r^x, k'_0, and d. Once the latter has been estimated from the coefficient of the linear term in p--call the estimate $\hat{d}$--the definitions of b_2 and b_3 give us two equations in two unknowns, r^x and k'_0. From the estimates $\hat{b}_2$ and $\hat{b}_3$ we may thus compute $\hat{r}^x$ and $\hat{k}'_0$. Mincer's results[72] imply values of 12.1% and 0.54, respectively, with d equal to 1.2%.

Although the estimate of k'_0 is well below unity (the value implied in Haley's theoretical model), Mincer considers it "rather high." He accepts without comment the estimate of d, though a lack of interpretation here may be somewhat misleading. If it is true, as argued in the previous section, that depreciation and growth are indistinguishable except in algebraic sign, then the coefficient labelled $\hat{d}$ must really measure (d - g) rather than d alone. Growth at an assumed rate of 2.5 - 3.0% would therefore mean depreciation at the rate of 3.7 - 4.2%. Johnson and Hebein,[73] with data able to distinguish growth and depreciation, encounter values of d in the range 1.0 - 3.4%. Haley's estimates[74] reach 4.3%. Thus Mincer's finding remains credible, even though considerably inflated by the suggested re-interpretation. Fortunately, we do not require a distinct estimate of d, but only the existing composite $\hat{d}$, in order to compute values for the other parameters, r^x and k'_0.

Turning to the quadratic specification (46a), we see that, unlike the exponential, it does not allow us to identify any of the parameters. The definitions of b_2 and b_3 represent two equations in four unknowns, r^x, k_0', T', and d. Mincer purports to eliminate one unknown (d) by expressing the model in net terms; that is, he ignores d and substitutes k(p) and T in place of k'(p) and I' in Equation (45).[75] This procedure raises no difficulty in the case of the integral, but in the case of the final, logarithmic term it appears invalid. The logarithmic term, it will be recalled, portrays the gap opened between measured and potential earnings on account of current investment in human-capital. This gap must surely depend upon gross rather than net investment. Mincer's procedure seems legitimate only for the special case in which depreciation equals zero. Then, gross and net investment are the same thing.

If one were prepared to assume zero depreciation, it would at first seem possible to identify the remaining parameters; for in this situation, measured and potential earnings reach an identical maximum where p = T = T'. One may locate maximum measured earnings by differentiating Equation (44a) with respect to p and setting the result equal to zero. The solution yields $p = T = -b_2/2b_3$. In this way, Mincer's published regression coefficients imply that when weeks worked are free to vary, earnings peak at 33.8 years of experience, and that when weeks worked are held constant, earnings peak at 37.8 years. Inserting these values for T in the equations defining b_2 and b_3 leads, however, to an inadmissible solution for r^x and k_0.

This outcome is by no means inexplicable. In the first place, it seems unlikely that depreciation is in fact equal to zero. Yet, if it were, one would have to recognize that under such circumstances, measured and potential earnings attain not so much a peak as a plateau, since in the absence of depreciation there is no reason for earnings (wage rates) to decline. It follows that unless T is actually very near retirement, the quadratic functional form may be inappropriate.

In practice, Mincer decides--arbitrarily it seems--to let T equal twenty years with weeks variable and thirty years with weeks held constant. Mysteriously, however, his published estimates-- $r^x = 6.3\%$ and $k_0 = 0.58$ in the first case, $r^x = 11.9\%$ and $k_0 = 0.42$ in the second--seem in arithmetic accord only if T were to equal 20.6 years and 33.1 years respectively.[76] In view of the theoretical problems just discussed, one cannot in any event put great store in the preceding results.

As predicted, the insertion of experience has a dramatic effect on the schooling coefficient. With postschool investment held constant in this parametric fashion, the estimated return to schooling increases from 7% in Mincer's Equation (S1) to about 11% in Equations (P1)-(G4). The exact specification of the investment profile has little bearing on the result.

There is, however, an even simpler method of holding post-school investment constant, and that is to consider only those individuals at a given stage of the life cycle. Mincer argues that the appropriate stage occurs at the point of overtaking.[77] At the overtaking year of experience ($\tilde{p}$), the individual earns, by definition, precisely the amount he would have received had he not engaged in any postschool investment. Hence, the earnings differentials observed within the overtaking set or cross-section are due entirely to differences in schooling. Rates of return computed from these earnings differentials will thus be free of bias.

We can obtain the "unbiased" estimates from the schooling model, provided we know the approximate period of overtaking. As explained in Chapter I, $\tilde{p} \leqq 1/r^x$. Mincer assumes: (a) that the preceding relationship holds with equality and (b) that $r^x = r^e$. Thus if r^e were equal to 12.5% $\tilde{p}$ would equal 8 years. After some experimentation Mincer settles on a cross section of individuals with 7 - 9 years of experience, producing Equations (V1)-(V4). Consistency demands that the rates of return estimated from these regressions equal approximately 12.5%. In Equation (V1), with weeks worked free to vary, r^e = 16.5%; in Equation (V2), with weeks worked held constant, r^e = 12.1%. The latter estimate is consequently the more pleasing of the two. However, both yield the hoped-for increase in the rate of return.

The weeks-constant estimate of r^e satisfies a further consistency requirement in that it comes close to Mincer's estimates of r^x. Had there been a large discrepancy, the definition of the overtaking set would have been suspect. At the same time, theory demands that $r^e = r^x$ at the margin; otherwise, the individual would not choose the level of schooling actually observed. Since Mincer assumes that r^e and r^x are constant, we must have equality as well in the estimated averages. We shall look for this consistency property in the results of Chapter III.

THE GENERAL MODEL

Though the empirical work of this study pertains solely to the special models of human-capital accumulation which we have already considered, it will be helpful in assessing and categorizing the present effort to examine, very briefly, the implementation of the "general model." At the level of theory, the general (Ben-Porath) model promises an integrated treatment of schooling and on-the-job training. However, when we come to implementation, this potential remains substantially unfulfilled. So far, researchers have been forced to apply the concepts of the model to homogeneous educational groupings, estimating distinct sets of parameter values in each case. What survives of generality must be found in the relatively wide class of postschool earnings profiles which the model can support.

The principal studies in the field are those of Ben-Porath, Heckman, Brown, Haley, and Moreh.[78] We have already noted in the preceding pages some of their quantitative results. Instead of merely assuming a convenient trajectory for postschool investment, this line of inquiry rests upon the deeper microeconomic foundation of a production function for human-capital. Not unexpectedly, therefore, the estimating equations turn out to be inherently nonlinear. The studies named utilize a variety of nonlinear methods. These differ chiefly in the parameters which the respective authors choose to specify rather than estimate. Thus Heckman fixes the discount rate; Brown, the rate of depreciation; Moreh, the production parameter (μ) and the age of retirement. Haley frees all the parameters but cannot identify the entire set. His estimating equation is by far the most complex of those surveyed.[79] The values which it can distinguish are generally plausible, and on this ground the Ben-Porath model derives support. The other studies turn up contradictions.

A notable feature of the preceding work is the small number of variables which it employs. Aside from the personal attributes (schooling and sex) which help in defining the various subsamples, only earnings (or their rate of change) and some variant of calendar time (either age or experience) take part in the calculations. The authors listed above all try to advance the basic model, not by capturing and inserting new information through the use of additional variables, but by estimating increasingly complex functional representations of the earnings profile. Despite the theoretical basis for this research, one is tempted to label it "curve fitting."

The problem resides, no doubt, in the practical limitations which beset nonlinear estimation procedures. These do not readily admit large data matrices. Because of the consequent need to restrict sample sizes, it is very difficult to treat general populations, which manifest considerable diversity. In small samples that are richly categorized, the cell frequencies often fall too low to give meaningful results. Even with a restricted sample, the researcher may not be able to include all the variables of interest.

The choices are therefore clear. One may settle for the rigourous estimation of a few hypothetical parameters, as in the case of Haley and the rest; or, one may sacrifice some degree of rigour, adopt an approximate specification for the human-capital investment profile, and pursue a broad investigation of the earnings structure. This study takes the latter approach.

APPENDIX IIA

TABLE 1

MINCER'S REGRESSION RESULTS [a]

Equations[b] (dependent variable: ln W)	R^2
Main Sample:[c]	
(S1) $7.58 + \underset{(43.8)}{.070s}$	.067
(P1) $6.20 + \underset{(72.3)}{.07s} + \underset{(75.5)}{.081p} - \underset{(55.8)}{.0012p^2}$	.285
(P2) $4.87 + \underset{(2.34)}{.255s} - \underset{(7.1)}{.0029s} - \underset{(31.8)}{.0043ps} + \underset{(63.7)}{.148p} - \underset{(66.2)}{.0018p^2}$	.309
(P3) $f(D_s) + \underset{(13.1)}{.068p} - \underset{(10.5)}{.0009p^2} + \underset{(119.7)}{1.207 \ln h}$	.525
(G1a) $7.43 + \underset{(77.6)}{.110s} - \underset{(102.3)}{1.6513e^{-.15p}}$	.313
(G1b) $7.52 + \underset{(74.3)}{.113s} - \underset{(101.4)}{1.52e^{-.10p}}$	.307
(G2a) $7.43 + \underset{(65.4)}{.108s} - \underset{(16.8)}{1.172e^{-.15p}} - \underset{(10.2)}{.32e^{-2(.15)p}} + \underset{(105.4)}{1.183 \ln h}$	.546
(G2b) $7.50 + \underset{(65.0)}{.111s} - \underset{(3.5)}{1.29e^{-.10p}} - \underset{(16.0)}{.162e^{-2(.10)p}} + \underset{(107.3)}{1.174 \ln h}$	.551
(G3) $f(D_{s,p}) + \underset{(108.1)}{1.142 \ln h}$	.557
(G4) $7.53 + \underset{(n.a.)}{.109s} - \underset{(n.a.)}{1.192e^{-.10p}} - \underset{(n.a.)}{.146e^{-2(.10)p}} - \underset{(2.4)}{.012p} + \underset{(n.a.)}{1.155 \ln h}$	.556
Overtaking Set:[d]	
(V1) $6.30 + \underset{(26.5)}{.165s}$	.328
(V2) $1.89 + \underset{(24.6)}{.121s} + \underset{(30.6)}{1.29 \ln h}$	.596

(Table 1 - continued)

Equations[b] (dependent variable: ln W)	R^2
(V3) $4.78 + .424s - .010s$ $(10.0) \quad (6.1)$	.347
(V4) $1.60 + .183s - .002s + 1.270 \ln h$ $(5.3) \quad (1.7) \quad (29.7)$	.602

Source: *Schooling, Experience, and Earnings,* p. 92, Table 5.1, and p. 53, Table 3.3.

[a]Figures in parentheses are t ratios, written in absolute terms.

[b]Original notation has been changed to conform with that employed in the current text. The symbol D refers to a vector of dummy variables for schooling and experience.

[c]28,678 observations on white, nonfarm, out-of-school males with experience not exceeding 40 years.

[d]2,124 observations on similar individuals with 7-9 years of experience.

APPENDIX IIB

BIASES IN THE EARNINGS FUNCTION DUE TO ERRORS IN THE MEASUREMENT OF EXPERIENCE

Let us suppose that the true earnings function is

$$Y = X\beta + u \ , \qquad \ldots . (IIB.1)$$

where

$$Y = \begin{bmatrix} \ln W_1 \\ \vdots \\ \ln W_n \end{bmatrix} \qquad X = \begin{bmatrix} s_1 & p_1 \\ \vdots & \vdots \\ s_n & p_n \end{bmatrix} \qquad \beta = \begin{bmatrix} \beta_s \\ \beta_p \end{bmatrix} \ .$$

As in the text, W stands for wages or earnings, s for schooling, and p for experience--all scaled here in deviations from their respective means. The disturbance vector u is assumed to have the classical properties

$$E(u) = 0 \ , \qquad E(uu') = \sigma_u^2 I \ , \qquad E(X'u) = 0 . \ldots . (IIB.2)$$

Suppose now that we observe $\tilde{Y} = Y$ and $\tilde{X} = X + V$ where V, the matrix of measurement errors, is given by

$$V = \begin{bmatrix} 0 & v_1 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ 0 & v_n \end{bmatrix}, \quad \text{whence} \quad \tilde{X} = \begin{bmatrix} s_1 & p_1 + v_1 \\ \cdot & \\ \cdot & \\ \cdot & \\ s_n & p_n + v_n \end{bmatrix}$$

Hence, p is the only variable measured with error.[1] We shall assume that $E(V'u) = 0$. It follows that

$$E(X'u) = E(\tilde{X}'u + V'u) = 0 \quad . \qquad \text{....(IIB.4)}$$

Substituting into (IIB.1), we obtain

$$\tilde{Y} = (\tilde{X} - v)\beta + u = \tilde{X}\beta + u - V\beta \quad . \qquad \text{....(IIB.5)}$$

Under these conditions, an ordinary least-squares regression of $\tilde{Y}$ on $\tilde{X}$ will yield the estimator $\hat{\beta}$, for which the expectation is

$$\begin{aligned} E[\hat{\beta}] &= E[(\tilde{X}'\tilde{X})^{-1}\tilde{X}'\tilde{Y}] \\ &= E[(\tilde{X}'\tilde{X})^{-1}\tilde{X}'(\tilde{X}\beta + u - V\beta)] \\ &= \beta - E[\tilde{X}'\tilde{X})^{-1}\tilde{X}'V\beta] \quad . \end{aligned} \qquad \text{....(IIB.6)}$$

Let us use $B = [B_s, B_p]'$ to represent the asymptotic bias in $\hat{\beta}$. Accordingly,

---

[1] Errors in Y merge with the components of u if we assume for them the same correlation properties. Therefore, nothing essential is lost by letting $\tilde{Y} = Y$.

$$B = \text{plim } [\hat{\beta} - \beta]$$

$$= \text{plim } [-(\tilde{X}'\tilde{X})^{-1}\tilde{X}'V\beta]$$

$$= -\text{ plim } \begin{pmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{pmatrix} \begin{pmatrix} s_1 & & s_n \\ p_1 + v_1 & \cdots & p_n + v_n \end{pmatrix} \begin{bmatrix} 0 & v_1 \\ 0 & v_n \end{bmatrix} \begin{pmatrix} \beta_s \\ \beta_p \end{pmatrix}$$

$$= -\text{ plim } \begin{pmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{pmatrix} \begin{pmatrix} \Sigma s_i v_i / n \\ (\Sigma p_i v_i + \Sigma v_i^2)/n \end{pmatrix} \cdot \beta_p \;, \qquad \text{. . (IIB.7)}$$

where the x_{ij} are elements of $(X'X)^{-1}/n$.

On the basis of arguments given in the text, we may hypothesize that:

$$\text{plim } \frac{\Sigma s_i v_i}{n} \leqq 0 \qquad \beta_p > 0$$

$$\text{plim } \frac{\Sigma p_i v_i}{n} > 0$$

$$\text{plim } \frac{\Sigma p_i p_i}{n} < 0 \quad . \qquad \text{. . . . (IIB.8)}$$

In addition, it is obvious that plim $\Sigma v_i^2/n > 0$. We assume that the preceding asymptotic variance and covariances converge to finite limits. Now, from (IIB.7),

$$B_s = -\text{ plim } \frac{1}{n} \left\{ x_{11} \Sigma s_i v_i + x_{12}(\Sigma p_i v_i + \Sigma v_i^2) \right\} \cdot \beta_p \quad ; \qquad \text{. . (IIB.9)}$$

$$B_p = -\operatorname{plim} \frac{1}{n}\left\{x_{12}\Sigma s_i v_i + x_{22}(\Sigma p_i v_i + \Sigma v_i^2)\right\} \cdot \beta_p \quad \ldots\ldots (\text{IIB.10})$$

To sign these expressions, we must investigate the elements of $(\tilde{X}'\tilde{X})^{-1}/n$.

Therefore, observe that

$$\frac{(X'X)^{-1}}{n} = \frac{1}{n|\tilde{X}'\tilde{X}|} \cdot \begin{bmatrix} \Sigma(p_i + v_i)^2/n & -(\Sigma s_i p_i + \Sigma s_i v_i)/n \\ -(\Sigma s_i p_i + \Sigma s_i v_i)/n & \Sigma s_i^2/n \end{bmatrix} .$$

$$\ldots (\text{IIB.11})$$

Since $(\tilde{X}'\tilde{X})$ is a positive definite variance-covariance matrix with a positive determinant, it follows with the help of our hypotheses that

$$x_{11}, x_{12} > 0 \quad \text{and} \quad x_{12} > 0 \quad .$$

We now have all the required information. From (IIB.9) and (IIB.10) it is apparent that if $\Sigma s_i v_i = 0$ asymptotically, both β_s and β_p will have a downward bias. A positive correlation between p and v makes this bias more severe. On the other hand, if $\Sigma s_i v_i < 0$, the bias in both coefficients is indeterminate, assuming we do not know the magnitudes of the correlations involved. Within the framework explored here,[2] the schooling coefficient β_s may have the upward bias suggested by Blinder *only* as a result of some negative correlation between s and v.

[2]A slightly more general model has been put forward by Maurice O. Levi, "Errors in the Variables Bias in the Presence of Correctly Measured Variables," Econometrica, XLI (September, 1973), 985-986. This derivation admits any number of independent variables but yields essentially the same results as encountered here.

It is of course well known that if more than one independent variable (in the present context, schooling) is measured with error, then no qualitative conclusions are possible. However, in the two-variable case, Theil has provided a helpful approximation formula,[3] which in our current notation reads as follows:

$$B_j \cong \frac{-1}{1-\rho^2}(\theta_j\beta_j - \rho\theta_k\beta_k) \qquad j \neq k = s, p, \ldots \text{(IIB.12)}$$

where ρ is the correlation coefficient linking s and p and θ_j is the ratio of the error variance in j to the variance of the true variable. Ceteris paribus, it would seem that errors in the measurement of schooling tend to lower both $\hat{\beta}_s$ and $\hat{\beta}_p$, since β_s and β_p are positive and ρ is negative.

---

[3]H. Theil, Economic Forecasts and Policy (Amsterdam: North-Holland Publishing Company, 1961), p. 329.

NOTES

CHAPTER 2

[1]Human Capital, p. 159.

[2]For convenient reference, all of Mincer's reported regressions have been reproduced in Appendix IIA, which follows this chapter. See Equation (S1).

[3]The best known examples are: Becker, Human Capital; Hansen, "Total and Private Rates of Return to Investment in Schooling"; Hanoch, "An Economic Analysis of Earnings and Schooling."

[4]The exceptions occur at very high and at very low levels of education. According to Hanoch, op. cit., marginal returns in the elementary grades sometimes exceed 100%, whereas, marginal returns to graduate education are 7% or less.

[5]Incomes of Canadians, p. 42, Table 5.9.

[6]Returns to Investment, p. 100, Table 5.14.

[7]In the following expression $\hat{r}_0^e$ and $\hat{r}_1^e$, are the estimated coefficients of s and s^2, respectively.

[8]See the Appendix IIA, Equations (V3) and (P2).

[9]Equation (V4).

[10]Schooling, Experience, and Earnings, p. 54.

[11]"Human Capital Theory," p. 838, n. 16.

[12]Appendix IIA, Equations (V2) and (V3). Note that these are marginal rates, the first having been assumed constant and therefore equal to the average. The mean level of schooling is given by Mincer as 12.2 years.

[13]In the simple case, with the rate of return assumed constant, $\bar{r}^e$ served to represent the average over years of schooling *and* over individuals populating the various schooling groups. With the rate of return allowed to vary, the average, as opposed to the marginal return, for individuals with s years of schooling is given by

$$\int_0^s (r_0^e + 2r_1^e t)\, dt \,/\, s \;=\; r_0^e + r_1^e s \quad ,$$

and the population mean is

$$\bar{r}^e \;=\; \int_0^\infty (r_0^e + r_1^e s)\, f(s)\, ds \quad ,$$

where f(s) is the proportion of individuals with less than s years of schooling.

[14]Haessel and Kuch, "Earnings in Canada," employ a three-step, nonlinear, iterative procedure to circumvent the heteroskedasticity problem. They do not report the extent to which the resulting maximum-likelihood estimates differ from those produced by ordinary least squares.

[15]Viewed in detail, the dependence of schooling on earnings may arise in several ways. As noted, earnings act on school attainment through the individual's investment response. If schooling is a normal consumption good as well as a repository of investment, individuals expecting (and later realizing) high earnings will make large "purchases." If the capital market is imperfect, initial earning capacity (embedded empirically in W_i) may constrain both consumption and investment. As pointed out in a slightly different context by G.S. Tolley and E. Olsen, "The Interdependence between Income and Education," Journal of Political Economy, LXXIX (May/June, 1971), 460-480, the preceding considerations apply not only to individuals but also to communities. Wealthy jurisdictions will spend more on education then poor ones, reinforcing individual tendencies.

[16]In the expressions $L_{(s)}^{dem}$ and $L_{(s)}^{sup}$, the subscript s in parentheses furnishes a reminder that we are really measuring different types or categories of labour on a single L axis. By including s in the argument lists of (33) and (34), we are thus able to treat compactly what is essentially a multimarket problem. Including mean earning or wage rates for the discrete labour types serves to emphasize the theoretical belief that quantities demanded and supplied depend on the full set of such rates. Alternatively, we could have inserted the continuous function $\bar{W} = \bar{W}(s)$. In this case, (33) and (34) become functionals. Note that in (33) demanders observe the true market averages. Imperfect knowledge on the part of demanders adds nothing of interest to the following analysis.

[17]At this point there is no need to be very specific about how expectations are formed. We need only be assured that expected wages respond to changes in actual market rates. In this static system we ignore whatever lags may be involved.

[18]We assume the existence of the multimarket equilibrium which this locus represents.

[19]A well known result in regression theory states that the product of the estimated slope coefficients must equal the square of the correlation coefficient between the two variables in question. Agreement in the estimate of r^e will thus occur only if the correlation between s and W is perfect.

[20]Making $r^e = r_0^e + r_1^e s$, with $r_1^e < 0$, does not change the present analysis, except that (39) below no longer provides an explicit solution for s_i.

[21]We assume that (35) captures individual expectations as well as the aggregate relationship originally portrayed.

[22]This conclusion is unaffected by making r^e depend on s in the manner proposed above. If we ignore the term $\ln[r_i/r^e]$ (either because it is small or because it vanishes when $r_i = r^e$), an explicit solution for s_i takes the form

$$s_i = \frac{-r_0^e \pm [(r_0^e)^2 - 4r_1^e\{\cdots + w_i \cdots\}]^{1/2}}{2r_1^e} .$$

Inspection will show the positive square root to be the relevant one. Accordingly, $d\, s_i/dw_i < 0$.

[23]Zvi Griliches, "Estimating the Returns to Schooling: Some Econometric Problems," Econometrica, XLV (January, 1977), 1-22. The degree of bias may be inferred by comparing Tables I and IV.

[24]Ibid., p. 13.

[25]See Tsuneo Ishikawa, "Family Structures and Family Values in the Theory of Income Distribution," Journal of Political Economy, LXXXIII (October, 1975), 987-1008.

[26]The latter include such things as unemployment and job vacancies, which may signal ensuing disequilibrium adjustment of wages and incomes. See The Market for College-Trained Manpower, pp. 8-10.

[27]*Ibid.*, pp. 59-60. In so doing, Freeman concurs with Theodore W. Schultz, who earlier suggested that uncertainty about future earnings was so great that individuals could not possibly refer to anything but current wages in determing their investments. See "The Rate of Return in Allocating Investment Resources to Education," *Journal of Human Resources*, II (Fall, 1967), 293-309, esp. pp. 303-305.

[28]See n. 3 above.

[29]Misspecification through the use of an incorrect functional form is not something about which we can speculate with any assurance.

[30]Strictly speaking, of course, we cannot determine how current wages might appear without specifying in full the underlying production function(s) and without ascertaining the regional and industrial pattern of output demand. However, the *direction* in which wages may appear to respond is in no way crucial to the present argument.

[31]We ignore, as usual, the finiteness correction

$$[1 - e^{0(r^e - g_0^*)T}]/[1 - e^{-(r^e - g_s^*)(T-s)}] \quad .$$

[32]Note the sign reversal in comparison with (16'). Because the latter is essentially an accounting formula, d enters there with a negative effect on earnings. By the same logic, g would appear with a positive sign. The equilibrating function is performed, if at all, by r'^e. In (41) r'^e is assumed fixed, and base-period (i.e., current) earnings make the necessary adjustment. Since these move in *compensatory* fashion, they rise with an increase in depreciation and fall with an increase in expected growth.

[33]Herman P. Miller, "Annual and Lifetime Incomes in Relation to Education," *American Economic Review*, L (December, 1960), 962-986.

[34]*Ibid.*, p. 73.

[35]See, in particular, Thomas Johnson, "Returns from Investment in Human Capital," *American Economic Review*, LX (September, 1970), 546-560; and Canada, Statistics Canada, *Economic Returns to Education in Canada*. The latter assumes a growth rate of 2.5%.

[36]Thomas Johnson and Frederick J. Hebein, "Investment in Human Capital and Personal Income, 1956-1966," American Economic Review, LXIV (September, 1974), 604-615, Table 1.

[37]"Estimation of Earnings Profiles," p. 1233, Table III. These and the preceding estimates appear to depend on how successful the authors are in accounting for endogenous growth through postschool investment.

[38]Supporters of the human-capital doctrine tend, naturally, to emphasize schooling and to minimize the role of all factors that are outside the individual's control. For a survey of the arguments see: F. Thomas Juster, "Introduction and Summary," in Education, Income, and Human Behavior, edited by F. Thomas Juster (New York, McGraw-Hill Book Company, Inc., 1975); or Sherwin Rosen "Human Capital: A Survey of Empirical Research" (Discussion Paper 76-82, Department of Economics, University of Rochester, 1976). Note especially Zvi Griliches and William M. Mason, "Education, Income, and Ability," Journal of Political Economy, LXXX (May/June, Supplement, 1972), S74-S103, and Samuel Bowles, "Schooling and Inequality from Generation to Generation," Journal of Political Economy, LXXX(May/June, Supplement, 1972), S219-S251.

[39]See the following: Herbert Gintis, "Education, Technology, and the Characteristics of Worker Productivity," American Economic Review, LXI (May, 1971), 266-279; Paul J. Taubman and Terence J. Wales, "Higher Education, Mental Ability, and Screenings," Journal of Political Economy, LXXXI (January/February, 1973), 28-55; Kenneth J. Arrow, "Higher Education as a Filter," Journal of Public Economics, II (July, 1973), 193-216; Richard Layard and George Psacharopoulos, "The Screening Hypothesis and the Returns to Education," Journal of Political Economy, LXXXII (September/October, 1974), 985-998; J.E. Stiglitz, "The Theory of Screening, Education, and the Distribution of Income; American Economic Review, LXV (June, 1975), 283-300; John G. Riley, "Information, Screening, and Human Capital," American Economic Review, LXVI (May, 1976), 254-260.

[40]On this account Becker deflated the rate-of-return estimated in Human Capital by 20%. Following Edward F. Denison, The Sources of Economic Growth and the Alternatives Before Us (New York: Committee for Economic Development, 1962), Bertram, The Contribution of Education to Economic Growth, applied a deflator of 40% to the Canadian data.

[41]See Griliches and Mason, op. cit.

[42]*Schooling, Experience, and Earnings*, p. 139.

[43]John C. Hause, "Earnings Profile: Ability and Schooling," *Journal of Political Economy*, LXXX (May/June, Supplement, 1972), S108-S138.

[44]See Gintis, *op. cit.*, and Finis Welch, "Human Capital Theory: Education, Discrimination, and Life Cycles," *American Economic Review*, LXV (May, 1975), 63-73.

[45]*Op. cit.*

[46]*Returns to Investment in University Training*, pp. 70-75.

[47]J. Behrman, Paul J. Taubman, and Terence J. Wales, "Controlling for and Measuring the Effects of Genetics and Family Environment in Equations for Schooling and Labor Market Success," in *Kinometrics: The Determinants of Socioeconomic Success Within and Between Families,* edited by Paul J. Taubman (Amsterdam: North-Holland Publishing Company, 1977).

[48]Paul J. Taubman and Terence J. Wales, "The Inadequacy of Cross-Section Age-Earnings Profiles When Ability is Not Held Constant," *Annals of Economic and Social Measurement*, I (July, 1972), 363-370.

[49]"Estimating the Returns to Schooling," pp. 4-6.

[50]*Op. cit.*, p. 67.

[51]Among the most prominent are: John T. Dunlop, "Productivity and Wage Structure," in *Income Employment and Public Policy* (New York: W.W. Norton & Co., Inc., 1948); Sumner, H. Slichter, "Notes on the Structure of Wages," *Review of Economics and Statistics*, XXXII (February, 1950), 80-91; Joseph Garbarino, "A Theory of Interindustry Wage Structure," *Quarterly Journal of Economics*, LXIV (May, 1950), 283-305; Leonard E. Weiss, "Concentration and Labor Earnings," *American Economic Review*, LVI (March, 1966), 96-117; Stanly H. Masters, "Wages and Plant Size: An Interindustry Analysis," *Review of Economics and Statistics*, LI (August, 1969), 341-345; Michael L. Wachter, "Relative Wage Equations for United States Manufacturing, 1947-1967; *Review of Economics and Statistics*, LII (November, 1970), 405-410; W. Hood and R.O. Rees, "Inter-Industry Wage Levels in United Kingdom Manufacturing," *Manchester School of Economics and Social Studies*, XLII (June, 1974), 171-183.

[52]*Op. cit.*

[53]C.T. Haworth and D.W. Rasmussen, "Human Capital and Inter-Industry Wages in Manufacturing," *Review of Economics and Statistics*, LIII (November, 1971), 376-380.

[54]"An Econometric Analysis of Earnings and Schooling," p. 312.

[55]See p. 66 above.

[56]*Op. cit.*

[57]See Greg J. Duncan, "Earnings Functions and Nonpecuniary Benefits," *Journal of Human Resources*, XI (Fall, 1976), 462-483; and Robert E. B. Lucas, "Hedonic Wage Equations and Psychic Wages in the Returns to Schooling," *American Economic Review*, LXVII (September, 1977), 549-558.

[58]See Appendix IIA.

[59]This explanation does not easily apply in the case of Mincer, who uses weeks rather than hours as the empirical counterpart of h.

[60]A further motive for inclusion, as we have seen, is to cancel variation in the rate of return to schooling.

[61]Schooling alone explains about 33% of the earnings variance. See Equations (V1) and (V2).

[62]See pp. 45-47.

[63]*Schooling, Experience, and Earnings*, p. 84.

[64]Haessel and Kuch, "Earnings in Canada," include women but subtract from experience a constant number of years for each child born. For other approaches see Jacob Mincer and Solomon W. Polachek, "Family Investment in Human Capita: Earnings of Women," *Journal of Political Economy*, LXXXII (March/April, Supplement), S76-S108; and Solomon W, Polachek, "Differences in Expected Post-School Investment as a Determinant of Market Wage Differentials," *International Economic Review*, XVI (June, 1975), 451-470.

[65]"On Dogmatism in Human Capital Theory," p. 14.

[66]Observe that individual measurement errors can always be written in the form $\bar{v}_p + v_{ip}$, where $\bar{v}_p$ represents the mean. If the latter exceeds zero, the mean level of experience will be inflated by a corresponding amount; but this distortion will not affect the value of any slope coefficients.

[67]"Male-Female Pay Differentials," Tables 1 and 2.

[68]"Human Capital: The Choice between Investment and Income," p. 937, Figure 5.

[69]Schooling, Experience, and Earnings, p. 84.

[70]Ibid., p. 93. The favoured values of β are 0.10 and 0.15. Mincer reports that: "While R^2 changes little in a wider internal, the partial repression coefficients are sensitive to the specification of β."

[71]The nearest comparison is probably between (P1) and (G1a) or (G1b), or between (P3) and (G2a) or (G2b).

[72]Schooling, Experience, and Earnings, p. 94.

[73]"Investment in Human Capital," p. 610, Table 1.

[74]"Estimation of the Earnings Profile," p. 1233, Table III.

[75]See Schooling, Experience, and Earnings, Chapter 4.

[76]Ibid., p. 94. Except for the last figure, these numerical results appear incorrect. The reader may wish to verify, using

$$k_0 = b_2T + 2b_3T^2 \qquad \text{and} \qquad r^x = b_2/k_0 - (1+k_0)/T_1$$

that the reported parameter estimates, together with the assumed values of T imply the following $r^x = 4.0\%$ and $k_0 = 0.66$ in the first case; $r^x = 11.5\%$ and $k_0 = 0.42$ in the second. Only in the second case is the discrepancy small enough to be attributed to rounding error.

[77]Ibid., pp. 47-49.

[78]Ben-Porath, "The Production of Human Capital Over Time"; Heckman, "Estimates of a Human Capital Production Function"; Brown, "A Model of Optimal Human Capital Accumulation"; Haley, "Estimation of the Earnings Profile"; Moreh, "Investment in Human Capital over Time."

[79]In fact, Haley's specification must surely be one of the most complex ever to appear in the econometric literature. See *op. cit.*, pp. 1228-1229, Equations (9) and (13).

CHAPTER III

THE EARNINGS FUNCTION: SINGLE-EQUATION ESTIMATES FOR CANADA

This chapter has two main objectives. The first is to present a series of estimates which reproduce with Canadian data the study of earnings functions carried out for the United States by Jacob Mincer. Though it is not everywhere prudent, given the multiple aims of the current study, or possible, given the data, to imitate Mincer's methods exactly, the procedures employed here yield results that are reasonably comparable. Some of the results, as we shall see, are virtually identical to Mincer's; others are strikingly different.

The second objective pursued in this chapter is to extend Mincer's investigation by adding to the earnings function variables which do not arise within a strict human-capital framework. Obviously, there are a number of factors besides schooling, experience, and weeks worked which influence the level of earnings. It is useful to isolate these factors statistically and to measure their relative importance, even though the associated hypotheses remain ad hoc. Omitting them could, if nothing else, bias the estimated coefficients of the human-capital variables. Whether or not any potential for bias actually exists, the expanded earnings functions appear to offer the best empirical standard against which to judge the performance of Mincer's undiluted human-capital specification. In the same way, these single-equation estimates serve as a basis of comparison for the system estimates reported in Chapter V.

The rest of the current chapter is divided into three sections. The second and third discuss, respectively, a Mincer-like set of human-capital regressions and a contrasting group of earnings function-estimates, expanded in the ways suggested earlier. Before we look at these empirical results, however, it is necessary to review the data and the methods which underlie them. Accordingly, the first section below describes in detail the principal data source used in compiling this study, the choices made in drawing the required sample, and the procedures followed in defining the many variables. Throughout this preliminary discussion, we shall take special note wherever an adopted procedure conflicts with one employed by Mincer.

THE DATA, THE SAMPLE, AND THE VARIABLES

The Principal Data Source

All the basic information used in this study originates with the 1971 Census of Canada. Except for one special tabulation, all of it comes, specifically, from the Public Use Sample, a vast set of individual records drawn from the Census Master File. The Public Use Sample (PUS) provides microdata on (1) individuals, (2) households, and (3) families resident in (a) the provinces or (b) the metropolitan areas of Toronto and Montreal. There are consequently six separate files, each furnished on magnetic tape.[1] This study employs the file on individuals resident in the provinces.

The Individual File, in common with the rest, is a one-in-one-hundred sample of the Canadian population. It is based on the Census long-form questionnaire, which was administered randomly to one-third of all households. A stratified random selection of one in every thirty-three and one-third such records provides the eventual one-in-one-hundred sample. The stratifying variables consist of age (three categories), sex (two categories), mother tongue (three categories), relation to head of household (three categories), and community type (three categories). The sample is thus representative of one hundred sixty-two distinct strata.

Each sample record supplies coded information on fifty-eight variables. The characteristics portrayed include among other things age and sex, place of residence, community type, the level of schooling and its geographic origin, the quantity, vintage, and type of vocational training, various aspects of family membership, the individual's language, citizenship, migration history, ethnic and religious background, labour-force status, industry and occupation, weeks and hours worked, total income, family income, income from wages and salaries, and income from self-employment--in short, a large array of economic and personal attributes. Needless to say, the PUS data do not supply any direct information on individual abilities or job experience. Of the fifty-eight characteristics available for study, twenty-nine contribute to the present research.

To preserve individual anonymity, the PUS tapes record much less detail for some characteristics than do the published Census reports. Industry, occupation, and place of residence are the variables chiefly affected. In the case of industry and occupation, the finer levels of disaggregation have merely been suppressed. There are twelve separate codes for industry and eighteen for occupation. In the case of residence, it was decided not to identify geographic areas with populations of less than 250,000. As a result, individuals living in Prince Edward Island, the Yukon, and the Northwest Territories were dropped from the sample. This omission, while unfortunate from the standpoint of completeness, could scarcely have had much effect on the overall regression estimates.

There is in general much similarity between the PUS data and the one-in-one-thousand sample Mincer obtains from the American census. However, in one important respect, the two bodies of information are quite incomparable. Mincer's sample pertains to 1959; the PUS data, to 1970. Hence, if we find some disparity in the regression estimates, it may be that Canada and the United States differ structurally; or it may be that the structures are identical but changing, and that we are simply measuring them at different points in time.

For the purpose of evaulating theoretically based arguments, it would be desirable, no doubt, to examine only contemporaneous comparisons. On other grounds, the problem of differing time periods does not seem especially significant. If Mincer's generalizations are "wrong" for Canada, it does not always matter whether they are wrong because they are outdated or because they fail to describe some unique features of the Canadian economy. It is chiefly important that such generalizations may prove misleading. Nevertheless, if the conclusions reached here contradict some of Mincer's, the theoretical appeal of the human-capital model is indeed diminished, since it is seen not to place binding restrictions on the data. Most researchers in the field would probably argue that the structures under consideration change rather slowly and that the greater part of any discrepancies uncovered must be the result of differences between the two countries. For this reason and for the others mentioned, the analysis presented below will not shrink from drawing the obvious comparisons, despite the incongruence in time periods.

The Sample

The PUS file selected as the principal data source contains information on just over 214,000 individuals. The first step in the research was to draw from this pool of records a working sample of manageable size and appropriate composition. With regard to sample size, the goal was to obtain 20,000 - 30,000 observations. This number is of the same order as that employed by Mincer and is well within the gross data-handling capabilities of the available computer software.

It is also large enough to provide adequate representation within all the designated population strata. With regard to sample composition, the problem was to exclude those individuals to whom the earnings model does not apply.

Since the model, as it stands, does not incorporate a theory of labour-force participation or unemployment, it cannot apply to individuals who report no work and, hence, zero earnings for the census year.[2] Negative earnings, which may arise through self-employment, are likewise inadmissible.[3] Individuals who did not work or suffered nonpositive earnings during 1970 were therefore excluded from the sample.

For essentially the same reason--inattention to time off work--the standard empirical model fails in attempting to explain the earnings of women. As we observed in Chapter II, the proxy designed to measure experience through the use of a single census cross section performs reasonably well only in the case of males. Females thus had to be eliminated from the sample.

Three other groups were also excluded: these in full-time attendance at a school or university, those employed in the public service (including the armed forces), and those whose industry of employment was "unspecified or undefined." The in-school population was excluded, first, because it is obvious that in this group individuals have not yet achieved the desired levels of education and, second, because any earnings they might report would likely be most atypical of what they could receive as full-time members of the labour force. Public servants were eliminated in order to focus as much as possible on individuals whose employers could be assumed to behave as profit maximizers.[4] Workers in unspecified or undefined industries were too few and too poorly characterized to warrant separate analysis; yet, they could not be combined satisfactorily with any other group. The best solution was therefore to ignore them.

A precise summary and technical statement of the sampling criteria may be found in Table 2. In light of the test for nonpositive employment incomes, the ones for zero weeks and zero hours are logically redundant but were nevertheless imposed to guard against inconsistency. All the listed criteria were applied in the given sequence to records from the PUS Individual File.

TABLE 2

SAMPLING CRITERIA

Individual Attribute	PUS Variable[a]	Codes Rejected[a]	Remarks
1. Sex	Sex	1	Excludes females.
2. Weeks worked in 1970	NUMWEEKS	0, 1	Excludes nonworkers, persons under 15 years.
3. Hours usually worked per week	USUALHRS	0	Excludes "not applicable."
4. Employment income	INCWAGES + INCSEL	sum ≤ 0	Excludes those with zero or negative earnings.
5. School attendance	ATTEND	1	Excludes full-time attenders (Part-time accepted).
6. Industry of employment	INDUST	00, 11, 12	Excludes nonworkers and persons under 15 years, workers in public administration and defence, workers in industries undefined.

[a]See Canada, Statistics Canada, Public Use Sample Tapes: User Documentation.

The following procedure was used to obtain the desired sample size. The beginning record--either the first or the second--was chosen at random,[5] and the indicated tests were applied to every *second* observation in the source file. In all, 107,010 records were scanned to create a working sample of 22,682 individuals. These numbers suggest the fraction of the total population (21.2%) to which the conclusions of this study apply.

Since the PUS file records are arranged initially in random order within provincial blocks, and since the proportion tested is very large, there is little reason to fear a biased or unrepresentative sampling, despite the lack of any explicit stratification in the selection procedure. Some feeling for the character and composition of the sample may be gained by looking at Tables 16-25, which form Appendix IIIA. These tables report the distribution of employment income, total income, and family income by size category, and the distribution of age, residence, and industry by level of education, showing in the last two cases both the number of individuals in each cell and their average earnings. Also included are distributions covering occupation, period of immigration to Canada, ethnicity, and religious affiliation. Mean earnings for the 22,682 individuals in the sample were $7,233, about 10% higher than the published statistic for all males 15 and over who worked in 1970.[6] Mean age was 39.8 years, and the mean level of schooling, 10.0 years.

The sample described in Appendix IIIA is "large" in the style of Mincer, statistically speaking, but differs somewhat in composition.

As we have previously noted, Mincer studies "white, nonfarm, non-student men."[7] The present research excludes women and full-time students but does not reject farm residents or nonwhites. Because of the desire to survey the Canadian population as fully as possible, and because of the data-processing overhead required to draw a second sample solely for comparative purposes, it was decided not to implement Mincer's first two criteria. Since it is a relatively simple matter to hold ethnic group and association with farming constant in the regression analysis, little is lost by adopting this procedure.[8] In general, it is not clear why the human-capital model should not apply to farmers and nonwhites. It may be that whites and nonwhites differ in ways that affect the model parameters, and it may be that farmers receive substantial nonmarket earnings or that they report as labour income part of the return on physical capital; however, it seems best to provide for such complexities through appropriate statistical techniques.

The present research does eliminate public servants and military personnel, whom Mincer apparently includes. If governments merely follow the lead of profit-maximizing firms in setting the wage structure (and if public-service unions strive to imitate private-sector bargains), one might argue that the human-capital model--or more precisely, these aspects of it which depend on profit-maximizing behaviour--could still apply. To have assumed such a "competitive" outcome would, though, have violated the spirit of the current study, which is to investigate the interplay of human-capital processes and market imperfection. It would seem a priori that this interplay is best observed in the private sector.

In addition to the four criteria already discussed, Mincer imposes two alternate restrictions, thus defining a *pair* of samples. One excludes individuals 65 and over; the other, individuals with more than 40 years of work experience. In fact, Mincer publishes results only for the latter.[9] He does not provide any explicit justification for the exclusions, but one might reason that the hypothesized experience profiles are unlikely to fit well at the upper end of the age scale.[10] In any event, it was decided not to implement either of Mincer's restrictions here.

Owing to the inclusion of farm residents and older workers, the current sample is probably somewhat more heterogeneous than the one Mincer chooses. The level of inequality is certainly greater. Taking the logarithm of earnings, we find that here its variance is 0.767. In the case of Mincer, it is 0.694 in the group aged under 65 and 0.668 in the group with 40 or fewer years of experience.[11] How much of the evident disparity is the result of differences in sample composition and how much, the result of intercountry comparison, is impossible to determine.

The Variables

This subsection defines all the regression variables used in the present study. For quick reference, Table 3 (below) introduces the symbolic name affixed to each, lists the PUS source variable, and offers a brief description. The ensuing text explains the construction of the most important variables in some detail, analysing the various choices which presented themselves.

TABLE 3

SUMMARY OF THE VARIABLES

Regression Variable	PUS Source Variable(s)	Description
AGE	AGE	Age.
ASQ	AGE	Age squared.
DF	USFAMINC, INCWAGES, INCSELF	Dummy: = 0 when INCFAM = 0; = 1 when INCFAM ≧ 0.
DI	INCTOTAL, INCWAGES, INCSELF	Dummy: = 0 when INCOTH = 0; = 1 when INCOTH ≧ 0.
ETH1-ETH7	USETHNIC	Ethnic or cultural group: 1 = British Isles*; 2 = Western European; 3 = Eastern European; 4 = Chinese and Japanese; 5 = Jewish; 6 = Native Indian; 7 = Negro, West Indian, other.
FAMSIZ	FAM SIZE	Number of persons in the individual's "census family" (= 1 in the case of a "nonfamily person").
GEO1-GEO6	GEO-CODE	Place of residence: 1 = Atlantic region; 2 = Quebec; 3 = Ontario*; 4 = Manitoba-Saskatchewan; 5 = Alberta; 6 = British Columbia.

Table 3 (continued)

Regression Variable	PUS Source Variable(s)	Description
HEAD	FAM-MEMB	Head of a census family; 0 = nonhead or nonfamily person; 1 = head.
IM1-IM4	PRDIMMIG	Period of immigration to Canada: 1 = before 1946; 2 = 1946-1965; 3 = 1966-1971; 4 = Canadian born*.
INC	INCWAGES, INCSELF	Income from wages and salaries and employment (= INCWAGES + INCSELF). In logs.
INCFAM	USFAMINC, INCWAGES, INCSELF	Family income in excess of INC (includes all property income and the earnings of other family members). In logs.
INCOTH	INCTOTAL, INCWAGES, INCSELF	Nonemployment income of the individual (= INCTOTAL - INCWAGES - INCSELF). In logs.
IND1-IND10	INDUST	Industry of employment: 1 = agriculture; 2 = forestry; 3 = fishing and trapping; 4 = mining and oil wells; 5 = manufacturing*; 6 = construction; 7 = transport, communications, utilities; 8 = trade; 9 = finance, insurance, real estate; 10 = community, business, and personal service.
LAN1-LAN4	OFF-LANG	Official language: 1 = English only*; 2 = French only; 3 = both; 4 = neither.
LENC1-LENC4	LENCRS	Length of vocational training; 1 = no training*; 2 = 3-5 months; 3 = 6 months-3 years; 4 = more than 3 years.
MAJ	MAJSINC	Major source of income: 0 = sources other than self-employment; 1 = self-employment (farm or nonfarm).
OC1-OC12	OCCUPAT	Occupation: 1 = managerial; 2 = natural and social sciences; 3 = teaching; 4 = medicine and health; 5 = clerical; 6 = sales; 7 = services*; 8 = farming and other primary; 9 = processing, fabrication, assembly, and repair; 10 = construction; 11 = transport operation; 12 = other (includes religion and the arts).

Table 3 (continued)

Regression Variable	PUS Source Variable(s)	Description
P	AGE, EDUCAT	Experience (= AGE - B', where B' = S + 5.67 when B > 15, and B' = 15 otherwise).
PSQ	AGE, EDUCAT	Experience squared.
PX	AGE, EDUCAT	Exp(βP), β = 0.05, 0.10, ···, 0.30.
P2X	AGE, EDUCAT	Exp(2βP), β = 0.05, 0.10, ···, 0.30.
REL1-REL4	US-RELIG	Religion: 1 = Protestant*; 2 = Roman Catholic and Orthodox; 3 = Jewish and other; 4 = none.
S	EDUCAT, AGE, GEO-CODE	Years of schooling (estimated). See text.
SCOST	EDUCAT, AGE, GEO-CODE	Years of schooling with positive opportunity cost (= S - 9 if S < 9; = 0 otherwise).
SPHG1-SPHG7	SCHOOL, PLCBIRTH	Place of highest grade in school (up to secondary level): 1 = Atlantic region: 2 = Quebec; 3 = Ontario*; 4 = Manitoba-Saskatchewan; 5 = Alberta; 6 = British Columbia; 7 = the Yukon and Northwest Territories or outside Canada. Defaults to place of birth for those with no schooling.
SSQ	EDUCAT, AGE GEO-CODE	Years of schooling squared.
TMARG	(See text)	1 - marginal tax rate (estimated). In logs.
TYPE	TYPE-71	Community type: 1 = urban, population 30,000 and over; 0 = urban, population under 30,000, plus rural, farm and nonfarm.
USMAR	USMARST	Marital status: 0 = single, widowed, divorced, separated; 1 = married.
WEEKS	NUMWEEKS	Weeks worked during 1970, divided by 50. In logs.
WTIME	NUWEEKS	Weeks in 1970 times usual hours per week, divided by 50·40 = 2000. In logs.
XINCFAMDF	-	Interaction: INCFAM·DF
XINCOTHDI	-	Interaction: INCOTH·DI

Table 3 (continued)

Regression Variable	PUS Source Variable(s)	Description
XPGEO1-XPGEO6	-	Interaction: P'GEO
XPIND1-XPIND10	-	Interaction: P'IND
XPOC1-XPOC12	-	Interaction: P'OC
XPSQGEO1-XPSQGEO6	-	Interaction: PSQ'GEO
XPSQIND1-XPSQIND10	-	Interaction: PSQ'IND
XPSQOC1-XPSQOC12	-	Interaction: PSQ'OC
XSGEO1-XSGEO6	-	Interaction: S'GEO
XSINS1-XSIND10	-	Interaction: S'IND
XSOC1-XSOC12	-	Interaction: S'OC
XSP	-	Interaction: S'P
ZINC	(see text)	TMARG + INC

*Denotes reference group of a dummy set.

The variables appearing in Table 3 may be sorted for further discussion into the following six categories:

1. Income variables: INC, MAJ, INCOTH, INCFAM, DI, DF, XINCOTHDI, XINCFAMDF, TMARG, ZINC;
2. Time-worked variables: WEEKS, WTIME;
3. Human-capital and life-cycle variables: S, SSQ, SPHG, P, PSQ, PX, P2X, XSP, AGE, ASQ, LENC;
4. Variables thought to represent immobilities and other market factors: GEO, TYPE, IND, OC, all interactions involving these attributes;
5. Family-status variables: HEAD, USMAR, FAMSIZ;
6. Personal-background variables: LAN, ETH, REL, IM.

We shall consider each group in turn.

1. *Income variables*. The principal dependent variable used in this study is INC, the sum of wages, salaries, and self-employment earnings, expressed in logarithms. Two problems arose in its construction. The first is one frequently encountered in working with income data: the highest incomes are grouped together in a single, open-ended class. Although the PUS source variables INCWAGES and INCSELF communicate actual dollar amounts rather than dollar ranges for most individuals, those reporting an income of $75,000 *or more* are shown as receiving *exactly* $75,000. This difficulty was met by assuming a Pareto distribution for the upper tail and computing, on that basis, the mean in the open-ended class.[12] Individuals were then assigned this level of income. In fact, however, the problem turned out to be insignificant, as INC--much less INCWAGES or INCSELF separately--exceeded $74,999 for only 18 observations, or 0.08% of the entire sample.

The other, more serious problem had to do with the composition of self-employment earnings. It is likely that amounts reported under this heading are a mixture of the returns to both human and nonhuman capital. Ideally, one would like to estimate the proportion attributable to nonhuman sources and subtract it in computing INC. Unfortunately, the available data (on unincorporated business) do not appear to warrant such an attempt. An alternative would have been simply to exclude individuals with positive (or large) self-employment earnings. This tactic would obviously have injected its own bias into the results, eliminating, for example, most individuals in the professions. As a compromise, it was decided to include self-employment earnings in the variable INC but to define, in addition, the independent dummy variable MAJ, which equals 1 when self-employment earnings are the major source of total income, and 0 otherwise. For individuals receiving only self-employment earnings, the use of MAJ is equivalent to assuming that the proportion of such earnings attributable to nonhuman capital is constant (though estimable and not specified in advance). However, since self-employment may affect equilibrium earnings in various ways,[13] we cannot impose any narrow theoretical interpretation on the coefficient of MAJ. Apart from the descriptive information to be gathered from this variable, its main purpose will be to counteract biases threatening other regression coefficients on account of the problem just discussed.

The income variables remaining in the list after INC and MAJ all contribute, in Chapters IV and V, to the empirical analysis of time worked. For completeness we shall nevertheless review their definitions here. INCOTH is a theoretical construct best understood as depicting the property or nonemployment income of the individual after personal income taxes. It was computed by subtracting from total income (PUS variable INCTOTAL) the sum of (a) estimated tax payments and (b) employment earnings multiplied by one minus the marginal tax rate (see below). As explained in Chapter IV, the result is used in mapping the individual's budget constraint. Alternatively, INCFAM measures all income of the family in excess of what the individual in question earns from employment. It was found by subtracting the two previously stated quantities (a) and (b) from total family income, as given by the PUS variable USFAMINC. Since the latter is in grouped form, class midpoints were used in this calculation.[14] Observe that (the antilog of) INCFAM equals (the antilog of) INCOTH plus both the property and employment incomes of other family members. However, INCFAM does not take into account other members' tax payments.

These definitions raise one complication: when, as sometimes happens, "other" incomes and own taxes equal zero, we cannot transform into logarithms. The solution in such instances was to let INCOTH or INCFAM equal some arbitrary value and define the interaction terms XINCOTHDI and XINCFAMDF. As explained in Table 3, the dummy variables DI and DF equal zero when the associated income variables equal zero; hence, so do XINCOTHDI and XINCFAMDF. Otherwise, XINCOTHDI = INCOTH, and XINCFAMDF = INCFAM. In practice, then, a dummy-interaction pair does the work of INCOTH or INCFAM, which never actually appear in any regression.

The last two variables related to income--ones also needed in the analysis of time worked--are TMARG and ZINC. As stated in the table, TMARG equals one minus the individual's estimated marginal tax rate (in logarithms). ZINC is simply TMARG + INC. Since the latter are both in logarithms, we have--once again in logarithms--the quantity (1 - marginal tax rate) × (employment earnings). This somewhat unorthodox construction stems from the analysis reviewed in Chapter IV.

Estimating TMARG meant, of course, simulating as carefully as possible the individual's personal income tax return. This task required certain assumptions and approximations. In the case of married family heads, it was assumed that the income of other family members (INCFAM - INCOTH) belonged solely to the spouse (here, necessarily, the wife) and that family size minus two measured the number of wholly dependent children.[15] In the case of nonmarried family heads, the number of potentially dependent children was assumed to equal family size minus one, with other income divided evenly among the subordinate individuals. Those who were not family heads were assumed to claim no dependents. Although the preceding five assumptions doubtless fail in many instances, they probably represent the great majority of family situations occurring in the present sample.

These assumptions, together with information on the 1970 tax structure,[16] would have been sufficient to determine total personal exemptions, except for one detail. The allowance for a dependent child varied in 1970, as it does currently, with the child's age. The present data source does not provide this information. Accordingly, an average claim ($341) was computed[17] and employed in all cases.

To arrive at taxable income, the simulation routine added to personal exemptions an average figure representing various common deductions which individuals are allowed. These involve registered pension fund and retirement savings plan contributions, medical expenses, charitable donations, and union or professional dues. Separate averages (of all such items combined) were computed in each of fourteen income classes.[18] The appropriate figure was then added to personal exemptions, as stated, and the result subtracted from total income (PUS variable INCTOTAL) to estimate taxable income.

The final step in the routine was to search through a table of effective marginal rates to find the one applying to the individual in questio. Since the combined federal and provincial rates vary across the country, it was necessary to take into account the individual's province of residence (PUS variable GEO-CODE).[19] A federal tax reduction prevailing in 1970 and special provisions relating to Quebec were also considered. The resulting estimate, labelled TMARG, is probably the best that can be inferred using census data. Though undoubtedly subject to error, it does not appear misleading in any systematic way.

2. *Time-worked variables.* WEEKS and WTIME are the two alternative measures of employment constructed here. They serve as independent variables in the earnings-function estimates reported in this chapter and as endogenous variables in the simultaneous-equation estimates to be presented later.

Let us first consider the definition of WEEKS. This variable is based on the number of weeks during which the individual worked, for however short a time, in 1970. The Census and, consequently, the PUS variable NUMWEEKS do not furnish much precision in this area, breaking down the fifty-two-week year into just five intervals (1-13, 14-26, 27-39, 40-48, 49-52). WEEKS was obtained by taking the five class midpoints,[20] dividing each by 50, and transforming into logarithms. Roughly speaking, therefore, WEEKS is measured in terms of years; more precisely, it is scaled so that the employment of "full-time" workers (49-52 weeks) equals unity. In view of the logarithmic transformation, this normalization affects only the constant term in the forthcoming regressions

The alternative employment variable WTIME takes into account both weeks and hours. It is the product (in logarithms) of weeks worked in 1970 and hours usually worked each week. This measure, or ones similar to it, have been used widely by economists and statisticians to estimate annual hours,[21] notwithstanding the acknowledged imprecision. The main problem afflicting WTIME stems from the hours component. In the Canadian census, hours are reported either for the job held in the week preceding enumeration day (July 1, 1971) or, in the case of persons then unemployed, for the job of longest duration held since January 1, 1970. One would obviously prefer an average of hours worked per week in 1970, if such a thing were practical. The Canadian definition, which stresses *usual* hours, is probably less objectionable than the American counterpart, which traditionally asks for hours worked "*last week*"; but both are clearly subject to transitory, short-run disturbances. Fortunately, it is not essential for purposes of this study to use WTIME in computing the hourly wage rate. This common procedure is one which places the most strain on the credibility of the variable.

Since the PUS source variable USUALHRS is again discontinuous--there are, to be exact, seven intervals--it was necessary to employ the class-midpoint approximation, as in the case of NUMWEEKS. In this instance, however, there was a final, open-ended class (50 or more hours/week) to deal with. Unhappily, there does not also exist a well-established theoretical distribution which one may apply in order to estimate the mean in this open-ended class. An arbitrary value of 54 hours/week was therefore assigned. The chosen figures were divided by 40 and transformed into logarithms, and the result for each individual was added to WEEKS in order to arrive at WTIME. The latter is consequently scaled in terms of a work year fixed at 2000 hours.

3. *Human-capital and life-cycle variables* The first human-capital measure we have to define is, of course, schooling. The PUS variable EDUCAT distinguishes twelve different levels: no schooling, less than grade 5, grades 5-8, grades 9-10, grade 11, grade 12, grade 13, 1-2 years university, 3-4 year (without degree), 3-4 years (with degree), 5 or more years (without degree), 5 or more years (with degree). To define the continuous regression variable S, we must translate each given level of education into an appropriate number of years. "No schooling" provides an obvious zero point for the scale, and it is natural to let grades 11, 12, and 13 equal 11, 12, and 13 years of instruction respectively.[22] The other eight levels demand a keener analysis.

In view of the emphasis accorded schooling by the present study, it was thought essential to measure this variable with as much accuracy as the Census itself would allow. Therefore, it was decided not to resort to the standard class-midpoint assumption in translating the PUS variable EDUCAT. Instead, special tabulations were obtained from Statistics Canada giving the number of out-of-school males at each single grade of public school or year of university, by age group and place of highest grade.[23] It was then possible to compute, for each schooling interval (except the last two), a mean value *conditional* upon age group and place of highest grade. These conditional means were used to estimate the schooling attainment of the individuals included in the sample. For most of these falling into the last two, open-ended classes, values of 17.5 and 18.5 years respectively were assigned.[24] The exception was for those schooled in Ontario, which maintains a thirteen-year system of public education. Here, the assumed figures were 18.5 and 19.5 years.

It is difficult to say how much the preceding refinements affect the subsequent regression estimates. Within the lower schooling intervals, which contain a large proportion of individuals, variation among the computed means was not insubstantial. In the second schooling interval (grades 1-4), the range was 2.72 - 3.66 years; in the third (grades 5-8), it was 6.47 - 7.67 years; and in the fourth (grades 9-10), 9.35 - 9.91 years. Variation within the narrower, postsecondary intervals was rather slight, but most of the computed values fell uniformly 0.10-0.20 years above or below the class midpoint, depending on the interval in question.

Though it is difficult to assess the effect of substituting conditional means for class midpoints, one may at least be confident that the schooling variable S will not suffer any contamination from age or region[25] as a result of the presentation of the data in grouped form.[26] Hence, the true impact of S will not be attributed to either of these other factors.

It is worth noting, finally, in connection with S that the source variable EDUCAT furnishes somewhat more detail than Mincer had at his disposal. Instead of the twelve schooling categories available here, he could consult only eight. It is not clear how Mincer dealt with the grouping problem.

Though later discussion will concentrate upon S, an alternative measure SCOST was defined in an effort to portray the number of school years with a positive (market) opportunity cost. On the assumption that Individuals cannot work in the market prior to age fifteen, SCOST was set equal to S-9 if S < 9 and equal to zero otherwise.

Besides stating the individual's level of education, the PUS data tell where the subject completed his last year of public school. This information permits the construction of a rough, though perhaps useful set of proxies for the *quality* of schooling. The dummy string SPHG--place of highest grade--was accordingly defined in the manner set out in Table 3. Note that the Yukon and Northwest Territories and "outside Canada" have been merged into one group--call it "outside Southern Canada"--and that place of highest grade defaults in the case of those with no schooling to place of birth, identically categorized. SPHG, together with AGE, fix unambiguously the individual's educational milieu at a particular stage of schooling and thus jointly stand in place of a quality index. Strictly speaking, however, we obtain a means of holding quality constant only for one year of study. As an overall measure of schooling quality, SPHG (plus AGE) will be inaccurate to the extent that individuals migrate interregionally during their years of public school. Moreover, SPHG has nothing to say about post-secondary education. In view of how S was constructed, using SPHG, redundancy may also be a problem.

Let us now consider experience. The basic variable P was computed in the manner described earlier--that is, by subtracting from age the sum of years schooling (S) plus age at school entry. However, no individual was credited with experience ostensibly gained before age 15.[27] Age at school entry was assumed to equal 5.67 years.[28] This value, an average, springs from two prior assumptions: (a) that birthdays are spread uniformly over the calendar year, and (b) that children begin school in September of the year during which they achieve age 6.

Notice that the special tabulations which assist in the construction of S also contribute to the estimation of P. With the mean level of schooling and age inversely correlated within schooling intervals, the standard procedure would have led to a modest overestimate of P for young individuals and to a similar underestimate for older ones. In the same way, P would have been underestimated for those schooled, and possibly still resident in, educationally deprived regions.

The variables derived from P--PSQ, PX, P2X, and XSP--require only brief comment. P and its square, PSQ, implement the quadratic functional form discussed in Chapter II. PX and P2X do the same for the exponential. The latter take on different values as the parameter β is iterated in steps of 0.05 from 0.05 to 0.30.[29] XSP is of course the experience-schooling interaction which appears in Mincer's work.

The last human-capital factor to note is LENC, a string of dummy variables representing the duration of any vocational course or apprenticeship undertaken by the individual (or if more than one, that of longest duration). Unfortunately, owing to a lack of detail in the PUS source data, it was impractical to attempt any decoding into time equivalents. Vocational training in the formal sense is not a factor given separate treatment by Mincer. Investigating its impact on earnings is therefore a matter of special interest, even though the data permit only the roughest sort of empirical analysis.

4. *Immobilities and other market factors*. If the market for skills were everywhere perfectly competitive, as human-capital theory presumes; if the adjustment to momentary disequilibrium were always rapid; and if the nonpecuniary returns to various jobs were unimportant--then it would be unnecessary, in attempting to explain individual earnings,

to look much beyond the human-capital variables already discussed. The sole aim of empirical research would be to produce refined estimates of the human-capital stock. Yet, it seems hardly prudent, when viewing the labour market, to assume *a priori* that immobilities and other imperfections, "momentary" disequilibria, and nonpecuniary factors will all be negligible. The acceptance or denial of such a proposition demands empirical inquiry.

We shall therefore consider a number of variables which one may interpret as standing for nonpecuniary differentials or market imperfection. The first of these is the dummy vector GEO, signifying place of residence (on enumeration day, July 1, 1971). If individuals are perfectly mobile and have no geographic preferences, the regression coefficients of GEO should all turn out insignificant. Note that GEO departs slightly from the standard five-region segmentation of Canada, distinguishing the relatively populous and industrially separate economy of Alberta from those of the other two Prairie provinces.

A related dummy variable TYPE denotes community size. Individuals in rural areas and those in "small" towns (population under 30,000) were grouped together (TYPE = 0) primarily in order to stress the earnings experience of those in large cities (TYPE = 1).

As promised in Chapter II, strings of dummy variables were also defined to represent industry and occupation. The construction of IND was a straightforward decoding of the PUS variable INDUST. It is nevertheless important to observe that the industry associated with each individual is the one which provided either the job held in the week prior to enumeration, or failing that, the job of longest duration held since January 1, 1970. There is thus no guarantee that reported 1970 earnings (INC) were derived wholly, or even partly, from employment in the reported industry. To the extent that individuals changed industries during the period under consideration, we must expect IND to contain some error. However, since the error is unlikely to be in any way systematic, its only effect should be to weaken the explanatory power of the industry variables. If these remain significant despite the error, the case against the human-capital variables as the sole determinant of earnings is strengthened all the more.

The same remarks apply to the vector of occupational dummies, OC--though as conceded in Chapter II, the case for including occupation in an equation with schooling already present is not so strong as that for including region or industry. With regard to the detailed specification of OC, it was found necessary to exercise some mild restraint in the number of variables defined. As a result, eighteen PUS categories were collapsed into twelve.

The need to economize on the number of variables arose principally on account of the desire to investigate the interaction of IND, OC, and GEO with the human-capital measures S, P, and PSQ. Even so, the number of interaction terms in this set reached seventy-five, not counting those pertaining to reference groups. For reasons of economy and for other reasons which will become clear when we examine the results of the next section, interactions involving the forms PX and P2X were not defined.

5. *Family-status variables.* These factors were included in some of the earnings equations primarily for descriptive purposes. Though one may conceive hypotheses in which they exert causal effect on earnings (perhaps via "reservation wages") or in which they serve as proxies for certain "ability" attributes, it would be a mistake, no doubt, to consider them wholly exogenous.

The first of these variables, HEAD, distinguishes those who head a "census family." The latter comprises either a husband, a wife, and any never-married children, or one parent and at least one never-married child, all living together. This nuclear aggregation was chosen for study in preference to the so-called "economic family," on which information was also provided.[30] "Head" always refers in the census definitions to the husband or parent (here, necessarily, the father) of any age. The second variable, USMAR, distinguishes married individuals. Those who are single, divorced, separated, or widowed--that is to say, those who report no *current* spouse--were grouped together in the reference category (USMAR = 0). The last variable, FAMSIZ, represents the number of persons in the census family, except that in the case of nonfamily persons, FAMSIZ equals one. Where the PUS source variable FAM-SIZE indicated "ten or more persons" (another open-ended class), FAMSIZ was set arbitrarily-- at eleven if USMAR equalled zero, and at twelve if USMAR equalled one. In effect, the number of children was assumed constant, on average, in the two cases.

6. *Personal-background variables*. These factors also play a descriptive role in the regression equations, though it is reasonable to treat them as exogenous. As in the case of family-status variables, hypotheses have been suggested linking them to earnings and employment. We shall not stop to consider such arguments here, but rather in the appropriate empirical sections which follow.

The definitions of LAN, ETH, REL, and IM are all relatively straightforward. LAN is based on official language instead of mother tongue (also available) because of the policy significance adhering to the former in Canada. With regard to ethnic group (ETH), twenty-one PUS categories were combined for purposes of this study into a more manageable seven. In the shortened description of Table 3, "Western European" includes French, Austrian, Finnish, German, Italian, Netherlands, and Scandinavian; "Eastern European" includes Czech, Hungarian, Polish, Russian, Slovak, and Ukrainian.[31] With regard to religious group (REL), the procedure was to distinguish Protestants, Catholic and Orthodox, non-Christians, and those professing no religion. Thirteen PUS categories were combined into four. Finally, with regard to period of immigration (IM), the rationale was to identify "early immigrants (before 1946), postwar immigrants" (1946-1965), and "recent immigrants" (1966-1971). "Canadian born" furnished the natural reference group. Thus ten PUS categories were again collapsed into four.

HUMAN-CAPITAL EARNINGS FUNCTIONS

In this section we shall treat only a few of the one hundred sixty-eight variables just defined. Replicating Mincer's orthodox human-capital approach, we shall see how his tightly specified earnings functions performed with the Canadian data.

These equations differ from one another, most fundamentally, in the way experience is held constant. As we observed in Chapter II, Mincer attacks this problem either by restricting the sample to one experience cross section (the overtaking set) or by postulating the form of the investment profile. In fact, Mincer tests two functional forms, the exponential and the quadratic.[32] We thus have three approaches to consider. The next three subsections deal with each one separately, in the order just stated.

Before we proceed to the results, one or two general comments are in order concerning the mechanics of estimation. Because the decoded raw data matrix had the intimidating dimensions 22,682 by 168, it would have been highly inefficient, if not impossible in practice, to process it in the usual manner, reading each observation into the computer and carrying out various preliminary calculations every time a new series of regressions was required. Fortunately, all of the statistical procedures contemplated in this study (including the three-stage least squares of Chapter V) could be performed knowing only the moment matrix of raw data. Actually, since the matrix is symmetric, only one triangle was needed.

A versatile and efficient regression programme known as RLS[33] was used to compute the moment matrix, which was then stored for easy access. In practice, the final matrix was itself built up in stages, by the simple process of matrix addition. The intermediate matrices provided distinct random subsamples of the large main sample. These were used for preliminary testing. Final estimates were then carried out for the full set of observations. This procedure tends to minimize the statistical dangers of hypothesis testing when the data are to be extensively "mined" by comparing a number of alternative specifications. All the estimates displayed here were obtained using RLS, which accepts moment matrices as input.

The Overtaking Set

As we observed in Chapter II, Mincer tends to favour an empirical definition of the overtaking set which includes individuals with 7-9 years of experience. In the present sample there turned out to be 1,238 individuals who met this criterion (specifically, $7.0 \leq P \leq 9.0$). Their mean years of schooling were 10.85--somewhat greater than for the full sample-- and the variance of logged earnings was 0.629--as expected, somewhat less.

Results for this group, corresponding to Mincer's Equations (V1)-(V4),[34] are displayed in Table 4. The simple regression of INC on S implies a return to schooling of 10.0%. This rate and the level of R^2 fall considerably short of the values obtained by Mincer. The addition of

WEEKS lowers the estimated return by about one quarter. This fraction presumably measures the return component which individuals receive indirectly, through increased employment rather than through higher wages. Note that, contrary to Mincer's findings, the coefficient of WEEKS does not depart significantly from one. Earnings are almost exactly proportional to weeks worked; by implication, wage rates do not depend on the volume of employment--not even through a mutual positive correlation of both factors with worker ability.

TABLE 4

ESTIMATES FOR THE OVERTAKING SET[a]

Equation Number	Equations[b] (dependent variable = INC)							R^2
(CV1)	0.5214 (6.88)	+	0.1001 S (14.9)					.152
(CV2)	1.003 (15.3)	+	0.0741 S (13.1)	+	0.9573 WEEKS (24.1)			.424
(CV3)	0.4609 (2.66)	+	0.1117 S (3.66)	+	0.0005 SSQ (0.39)			.152
(CV4)	1.188 (8.16)	+	0.0392 S (1.55)	+	0.0015 SSQ (1.42)	+	0.9617 WEEKS (24.2)	.425

[a]1238 observations on individuals with 7-9 years of experience.

[b]Figures in parentheses are t ratios, written in absolute terms.

Schooling squared (SSQ) does not achieve significance whether or not WEEKS is included. The rate of return appears to be constant even when employment is allowed to vary. Thus Mincer's argument on this point turns out to be irrelevant, at least for the present group. However, looking at Equation (CV4),[35] where even S is insignificant, one begins to suspect that the quadratic functional form may be inappropriate in the Canadian setting. As we have noted, direct estimates for Canada have previously shown a somewhat irregular (nonmonotonic) pattern in the rates of return to schooling,[36] rather than the nearly continuous schedule of decline familiar in United States studies.

On the whole, Equations (CV1)-(CV4) do not seem especially favourable to the use of the overtaking concept. Except in (CV3), the implied rates of return are not consistent with the assumed length of the overtaking period (recall that if costs are constant, $\tilde{p} = 1/r^x$). One must bear in mind, however, that the length assumption, which defines empirically the overtaking set, was simply copied from the work of Mincer. If rates of return are lower in Canada than in the United States, a somewhat longer period of overtaking might have given better results. Since the search for a new empirical definition appears methodologically dubious, we shall not pursue this problem here. Instead, we shall turn to the full sample of individuals and to parametric methods of holding experience constant.

Exponential Experience Profiles

Besides holding experience constant so that one may estimate the return to schooling in an unbiased manner, the exponential form of the experience profile should allow one to estimate the initial propensity to undertake postschool investment (k_0), the typical net postschool rate of return (r^x), and even the rate at which human capital depreciates (d). From Equation (46b) in Chapter II it follows that

$$k_0 = (-2b_3)^{\frac{1}{2}} \quad \text{and} \quad r^x = \beta[(b_2/k_0) + 1] \quad ,$$

where b_2 and b_3 are, in the current notation, the coefficients of PX and P2X respectively. The coefficient of P, when that variable enters the regression along with PX and P2X, furnishes the estimate of d.

To be admissible, the implied value of k_0 must fall within the closed unit interval; that of r^x must surely be nonnegative (otherwise no one would think of investing). The preceding requirements place certain reasonableness restrictions upon b_2 and b_3, namely:

$$b_2 \leqq -k_0 \quad \text{and} \quad -1/2 \leqq b_3 \leqq 0 \quad .$$

If these conditions are not met simultaneously, the model fails.

The outcome of experiments with the exponential form appears in Table 5. These results, obtained by iterating for different values of β in the same way as Mincer, are not very encouraging. As β increases,

b_2 declines and b_3 rises, each monotonically. None of the specific values tried for β produces coefficients which meet the reasonableness requirements. Viewing Equations (CG2), one might expect, on the basis of monotonicity, to encounter reasonable coefficients when β is in the 0.15-0.20 range. Unfortunately, there does not appear to be much hope of refining this estimate. As was reportedly the case with Mincer's sample, the value of R^2 does not change significantly within the plausible range of β. It is not clear what other criterion one could possibly use. Mincer, of course, relies on the plausibility of the coefficients themselves, or equivalently, upon r^x and k_0; but this course is not open here. One could presumably search over values of β in the 0.15-0.20 range and obtain plausible figures for r^x and k^0, but one could not then claim to have "estimated" these parameters. In view of how sensitive b_2 and b_3 seem to be, a great many pairs of values would likely be found acceptable. One's general conclusion must be that the exponential form is not a satisfactory device for estimating the investment parameters in the case of Canadian males.

The other results presented in Table 5 reinforce this inference. In Equations (CG4) and (CG5), the admissible values of β must be somewhat less than 0.05. It is difficult to believe that an optimal plan would dictate such a low rate of decline (under 5%) in the net propensity to invest, given the length of the average working lifespan. In (CG4) the coefficients of P, interpreted as rates of depreciation, are not alone implausible; but in light of the suspicion surrounding, first, the value of β and, second, the functional form, they cannot be taken very

TABLE 5

FULL-SAMPLE[a] ESTIMATES USING EXPONENTIAL EXPERIENCE PROFILES

Equation No.	β	Coefficients[b] of S	SSQ	P	PX	P2X	WEEKS	R^2
(CG2)	.05	0.0745 (52.9)	-	-	2.160 (27.9)	-3.026 (37.7)	0.8411 (82.4)	.406
	.10	0.0768 (53.5)	-	-	0.3274 (5.36)	-1.604 (20.7)	0.8531 (82.9)	.397
	.15	0.0782 (54.6)	-	-	-0.6097 (9.72)	-0.6539 (7.71)	0.8587 (83.4)	.395
	.20	0.0784 (55.2)	-	-	-1.287 (19.2)	0.1099 (1.19)	0.8588 (83.5)	.396
	.25	0.0780 (55.2)	-	-	-1.814 (25.1)	0.7270 (7.30)	0.8573 (83.4)	.397
	.30	0.0772 (54.8)	-	-	-2.238 (28.9)	1.228 (11.6)	0.8561 (83.2)	.396
(CG4)	.05	0.0732 (51.3)	-	-0.0309 (13.9)	-1.572 (5.61)	-0.8364 (4.72)	0.8405 (82.6)	.411
	.10	0.0734 (51.3)		-0.0204 (22.6)	-2.080 (16.9)	0.2621 (2.33)	0.8398 (82.4)	.410
	.30	0.0730 (50.7)	-	-0.0057 (12.8)	-2.950 (31.0)	1.885 (16.0)	0.8466 (82.4)	.401
(CG5)	.05	0.0082 (1.51)	0.0030 (12.3)	-0.0316 (14.2)	-1.565 (5.61)	-0.8639 (4.90)	0.8430 (83.2)	.415
	.10	0.0086 (1.58)	0.0030 (12.3)	-0.0209 (23.2)	-2.066 (16.9)	0.2261 (2.01)	0.8421 (82.8)	.414
	.30	0.0098 (1.78)	0.0029 (11.9)	-0.0062 (13.9)	-2.946 (31.0)	0.844 (15.7)	0.8488 (82.8)	.404

[a]22,682 observations

[b]The dependent variable is INC. Figures in parentheses are t ratios, written in absolute terms. Constants, though present in all the regressions, are not shown.

seriously. Equations (CG5) depart slightly from Mincer in adding SSQ. Here, in contrast to (CV4), the term is significant, though S itself is not. The positive coefficient implies that r^e increases with the level of schooling--by about 0.6% for each additional year. In view of Podoluk's results[37] from the 1961 census, this finding is not a complete surprise, though again it is at variance with United States experience. Here, the indicated return at the mean year of schooling is just under 7%.

Quadratic Experience Profiles

Estimates obtained using quadratic experience profiles are shown in Table 6. These results are no more helpful in attempting to evaluate r^x and k_0 than are the ones derived using the exponential form, but they are perhaps easier to interpret from a purely descriptive standpoint.

Before we examine what little the estimates have to offer concerning the investment parameters, let us look at various other, more transparent implications. Note first of all the schooling regression (CS1), inserted in Table 6 for purposes of comparison. As it turns out, the schooling coefficient, when rounded, precisely matches that of Mincer. On the basis of R^2, schooling may be said to explain 7.3% of (log) earnings variance--just a little more than in Mincer's sample.

The addition of the experience term in (CP1) causes the schooling coefficient to rise, as expected--though not quite so markedly as in Mincer's (P1). Differentiating with respect to P (remembering that PSQ = P^2) and setting the result equal to zero show that earnings reach a peak

TABLE 6

FULL-SAMPLE[a] ESTIMATES USING QUADRATIC EXPERIENCE PROFILES

Ind. Variable	Equations[b] (dependent variable = INC)						
	(CS1)	(CP1)	(CP2)	(CP3)	(CP4)	(CP5)	(CP6)
Constant	.9906 (57.0)	-.0714 (2.86)	-.3663 (5.51)	.5397 (23.6)	.5809 (9.85)	.3944 (17.2)	-.7484 (17.2)
S	.0695 (42.4)	.0891 (54.8)	.1009 (11.3)	.0715 (49.9)	.0393 (5.04)	.0775 (53.5)	.0624 (46.0)
SSQ	-	-	.0009 (2.91)	-	.0022 (7.92)	-	-
P	-	.0829 (63.3)	.1029 (44.3)	.0583 (49.7)	.0683 (33.1)	.0572 (47.5)	-
PSQ	-	-.0014 (58.8)	-.0016 (58.0)	-.0010 (46.5)	-.0011 (45.0)	-.0010 (42.8)	-
XSP	-	-	-.0014 (10.3)	-	-.0007 (5.72)	-	-
AGE	-	-	-	-	-	-	.0983 (49.6)
ASQ	-	-	-	-	-	-	-0011 (47.2)
WEEKS	-	-	-	.8629 (85.3)	.8615 (85.3)	-	.8576 (84.8)
WTIME	-	-	-	-	-	.6589 (78.7)	-
R^2	.073	.213	.220	.405	.409	.382	.407

[a]22,682 observations.

[b]Figures in parentheses are t ratios, written in absolute terms.

at 29.6 years of experience. Holding weeks constant, in (CP3), lowers the estimated rate of return from 8.9% to 7.2%--that is, by about one fifth--but leaves peak earnings, at 29.2 years, little changed.

The insertion of SSQ and XSP, in Equations (CP2) and (CP4), helps to delineate further the shapes of the earnings profiles. As before with the Canadian data, the coefficient of SSQ is positive and significant, though admittedly rather small in the first case. Holding weeks constant does not eliminate the apparent rise in the rate of return but, in fact, seems to strengthen it.

Turning to XSP, we find that its coefficient is significantly negative. As Mincer points out, this result implies that experience profiles for the various levels of schooling tend to converge over the life cycle, since earnings rise less (or decline more) with experience at high levels of schooling than at low levels. The degree of convergence indicated here is nevertheless relatively small in comparison with that observed by Mincer.

When we take both SSQ and XSP into account, the implied rate of return to schooling for individuals with mean levels of schooling and experience (10.03 and 23.14 years respectively) turns out to be 8.7% with weeks variable and 6.7% with weeks held constant. For mean-schooled individuals, measured earnings peak at just under 28 years of experience in both cases. Differentiating the expression for the peak-earnings year with respect to S shows that an additional year of schooling hastens the peak by 0.3-0.4 years in terms of experience. In terms of age, the peak is therefore postponed by 0.6-0.7 years.

Replacing the quadratic in experience with a quadratic in age reveals in (CP6) that (weeks-constant) earnings peak, on average, at 44.7 years of age. At normal retirement, earnings will have receded by almost 20%, according to the estimates. The age quadratic fits the Canadian data just as well as, if not better than, the experience quadratic; but in the former case, the implied rate of return to schooling is lower and, perhaps, negatively biased.

Coefficients of the employment variables, representing elasticities, are significantly *less* than one throughout Table 6. This finding contrasts sharply with that of Mincer, who observed elasticities in the neighbourhood of 1.2. It is also at variance with the outcome in the overtaking set, for which the measured elasticities are not significantly different from one. The indication is that low wages and high levels of employment go together. This seems especially to be the case when we consider hours (WTIME) rather than weeks in (CP5). The implied elasticity drops from 0.86 to 0.66. The fit is slightly weaker than in (CP3), reflecting perhaps the errors to which WTIME is subjec. "Errors in variables" may indeed have some part in depressing the coefficients of *both* WEEKS and WTIME. However, it should be noted that Mincer's employment variable, with which we are making comparison, suffers the same shortcoming.

One may of course rationalize in various ways the apparent inelasticity of earnings with respect to employment. A backward-bending supply curve of labour would explain this result, especially if one assumes perfect competition. Workers confined to low-wage jobs may very well seek long hours or "moonlight" in order to reach equilibrium. In an environment of discrete choices, some workers may have such a strong taste for income that they eschew high-wage jobs with standard, inflexible weeks and hours in favour of low-wage jobs with weeks and hours unconstrained. The latter may occur even though individual supply curves are positively sloped. The trouble with both these arguments is that they require us to postulate radically different preferences, or distributions of preferences, among the Canadian and American work forces.

A superior explanation may therefore lie in the pronounced seasonality of economic activity in Canada. If seasonal workers are involuntarily unemployed during part of the year (or if they are simply earnings maximizers), they will demand, and in competitive equilibrium receive, high wages as a compensation for low hours. Despite the plausibility of this argument, it is probably unwise to speculate very far on the basis of the present single-equation estimates, which may be biased, and which doubtlessly entangle labour-supply, labour-demand, and investment responses. We shall take up the elasticity question again in light of the simultaneous-equation estimates reported in Chapter V.

It remains in this section to explore briefly what the present estimates imply concerning the investment parameters. From Equation (46a) and the accompanying definitions it follows that

$$k_0' = b_2 T' + 2b_3 T'^2 + d \cdot T'$$

and

$$r^x = b_2/k_0' - (1 + k_0')/T' + d/b_2 \quad ,$$

where b_2 and b_3 are the coefficients of P and PSQ respectively. Since it is not possible to identify the four unknowns (r^x, k_0', T', and d) using only the preceding pair of expressions, we must be content to examine a range of numerical combinations in order to see where the most plausible values lie. Table 7 shows, in the weeks-variable case, the values of r^x and k_0' which arise in connection with certain specified values of T' and d. Because one may wish to interpret the latter as the difference between depreciation and expected growth, some nonpositive values have been included for trial.

As much as anything, Table 7 seems to emphasize the inadequacy of the present technique for measuring the rate of return to postschool investment. If one is prepared to assume the validity of the model, then it is possible to rule out "large" values of T' and d; but there is little else that one may say. Over the six admissible cases--those in which, say, $0\% < r^x < 30\%$ and $0 < k_0' < 1$ --r^x ranges from 3.9% to 20.2%. The r^x-k_0' pair corresponding to $T' = 20$ and $d = 0$ is perhaps worthy of special note, since it is the combination implied by Mincer's assumptions. The values obtained here are similar to the ones Mincer reports; but as the table demonstrates, they are too sensitive to the assumptions concerning T' and d to warrant much confidence.

TABLE 7

VALUES OF r^x AND k'_0 CONSISTENT WITH SPECIFIED VALUES OF T' AND d (WEEKS-VARIABLE CASE)[a]

d		T' (years)		
		20	25	30
-.01	r^x =	14.8%	103.4%	-27.1%
	k'_0 =	.34	.08	-.33
0.00	r^x =	7.7%	20.2%	-273.2%
	k'_0 =	.54	.33	-.03
.01	r^x =	3.9%	9.9%	30.2%
	k'_0	.74	.58	.27
.02	r^x =	4.3%	10.6%	30.0%
	k'_0 =	1.14	.83	.57
.03	r^x =	4.8%	10.9%	30.0%
	k'_0	1.34	1.08	.87
.04	r^x =	5.1%	11.2%	29.9%
	k'_0 =	1.54	1.33	1.17

[a]See Table 6, Equation (CP1), in which b_2 = 0.083 and b_3 = -0.0014.

EXPANDED EARNINGS-FUNCTION ESTIMATES

Having considered the strict human-capital specification, we may now view the results obtained by expanding the earnings functions to include variables typically ignored by human-capital theorists. We shall pay particular attention to any changes which occur in the schooling coefficient as new variables are added. More generally, we shall be able to assess the relative importance of human-capital and other factors in determining the employment incomes of Canadians.

To begin the analysis, we must choose one of the human-capital earnings functions as a standard of comparison. The quadratic Equation (CP5) seems best suited for this purpose. It is simple to estimate and to interpret, and its functional form is by far the most widespread in the literature. Though (CP3), containing WEEKS, fits slightly better, statistical concerns arising later in connection with the system estimates of Chapter V favour the use of WTIME as the employment variable. Hence, (CP5) is to be preferred. We shall not ignore, however, the variables SSQ and XSP, which are missing from it. These terms will ultimately be included in the expanded regressions.

The latter are displayed and discussed in the first subsection below. The second deals with a particular version of the so-called "interactions model."

The Impact of Previously Omitted Variables

Earlier in this chapter, variables which might be thought to influence employment earnings were grouped under several headings. Restated here for convenience, they are: (1) human-capital and life-cycle variables, (2) variables thought to represent immobilities and other market imperfections, (3) family-status variables, (4) personal-background variables. The text and the tables which follow review each set of factors in turn. Further divisions examine an alternative to the initial specification, analyse the occupational dimension of employment earnings, and present a brief summary.

It must be noted, to begin, that the order in which variables enter succeeding regressions may have an effect on the interpretation of results. Because here, and in general, the independent variables of concern are correlated with one another, there will always be some area of indeterminacy in the assignment of explanatory significance. The amount by which a particular variable increases the level of R^2 is one estimate of its importance, but only a conditional estimate for the set of regressors included by prior selection. The order of selection established here follows principally from the emphasis given by this study to the variables in groups (1) and (2) above, we shall devote special attention to the indeterminacy or variance-attribution problem as it affects the preceding factors.

1. *Human-capital and life-cycle variables.* The main factors in this group which do not appear in the orthodox specifications are LENC and SPHG. With regard to the former, Table 8 shows that brief vocational courses (LENC1) have no significant effect upon earnings. Programmes of intermediate length (LENC2) have a modest effect at best (see also Table 9). However, long vocational programmes, which one might guess consist mainly of classical apprenticeships, add as much as 18% to the level of earnings (see the coefficient of LENC3 in Equation (CP7)).[38] Holding additional variables constant nevertheless reduces this apparent premium considerably. Vocational preparation is evidently correlated to a significant degree with both place of residence (GEO) and industry (IND), especially the latter. At a minimum (in Equation (CP13), Table 10), the apparent earnings premium associated with LENC4 falls to 8.0%.

As discussed earlier, SPHG (place of highest grade) may be considered a proxy variable for schooling quality. Not surprisingly, however, SPHG and GEO (place of current residence) turn out to be closely correlated. When both are entered in the same regression, some coefficients of GEO survive the ensuing multicollinearity; but those of SPHG become uniformly insignificant.[39] SPHG on its own does not match the performance of GEO under identical circumstances. Prelininary tests supporting these observations may be found along with other, miscellaneous regressions in Appendix IIIB. Further work utilizing SPHG was not attempted.

TABLE 8

REGRESSION ESTIMATES[a] OF THE EXPANDED EARNINGS FUNCTION, I

Ind. Variable	Equations[b] (Dependent variable = INC)					
	(CP7)		(CP8)		(CP9)	
Constant	.3977	(17.4)	.6323	(25.8)	.6616	(27.0)
S	.0763	(52.5)	.0688	(47.0)	.0705	(47.5)
P	.0564	(46.8)	.0548	(46.3)	.0525	(45.7)
PSQ	-.0009	(42.0)	-.0009	(41.5)	-.0008	(39.4)
WTIME	.6577	(78.7)	.6569	(80.0)	.6804	(84.7)
LENC2	-.0014	(0.05)	-.0011	(0.04)	-.0086	(0.31)
LENC3	.0363	(2.39)	.0260	(1.75)	.0115	(0.80)
LENC4	.1782	(9.42)	.1460	(7.85)	.0998	(5.53)
GEO1	-	-	-.1770	(9.87)	-.2085	(11.9)
GEO2	-	-	-.0604	(5.33)	-.0604	(5.51)
GEO4	-	-	-.2601	(15.2)	-.1422	(8.43)
GEO5	-	-	-.1058	(5.91)	-.0480	(2.75)
GEO6	-	-	.0491	(3.11)	.0306	(2.00)
TYPE	-	-	.1987	(21.0)	.1151	(11.9)
IND1	-	-	-	-	-.7047	(31.1)
IND2	-	-	-	-	-.0111	(0.32)
IND3	-	-	-	-	-.4090	(6.89)
IND4	-	-	-	-	.1943	(6.99)
IND6	-	-	-	-	.0557	(3.52)
IND7	-	-	-	-	.0365	(2.46)
IND8	-	-	-	-	-.1573	(11.5)
IND9	-	-	-	-	.0099	(0.42)
IND10	-	-	-	-	-.1282	(9.42)
MAJ	-	-	-	-	-.0477	(2.86)
R^2	.385		.409		.452	

[a]Main sample, 22,682 observations

[b]The first figure in each set is a regression coefficient; the second, in parenthesis, is the corresponding t ratio, written in absolute terms

Also relegated to the appendix is an illustrative equation employing SCOST in place of S. Recall that SCOST counts only those years of schooling registered after about age fifteen. It does so on the speculation that early school attendance may entail no opportunity cost and thus should not be presumed costly in deriving the model. As one might expect, especially in view of the results concerning SSQ, the rates of return implied for SCOST exceed those for S, the addition being about 1.5 percentage points. As one might also expect, SCOST does not yield as high an R^2 as S. Though differences in schooling at the low end of the scale may not reflect investment decisions, such differences are evidently recognized and rewarded in the market, either because schooling in the range under discussion enhances productivity or because it serves as a proxy for ability and background characteristics which we are otherwise unable to measure. Accordingly, S would appear to be the variable of choice in the analysis of earnings determination and distribution, even though its truncated variant SCOST might possibly give better rate-of-return estimates. Since replacing S with SCOST had little effect on the coefficients of *other* variables, we shall not pursue further experiments with the latter but will instead concentrate on S in order to present results of maximum comparative interest.

2. *Variables thought to represent immobilities and other market imperfections.* Prime candidates under this heading are GEO and IND. These are added sequentially, along with TYPE and MAJ, in

Equations (CP8) and (CP9), Table 8. As explained previously, the coefficients measure percentage differences in earnings relative to the chosen reference group. In the case of (CP9), the reference group consists of nonmetropolitan Ontario residents without formal vocational training employed as wage-earners in manufacturing.

It turns out that all the coefficients of GEO and TYPE are significant at the 0.05 level or better; indeed, all but one are significant at the 0.01 level. The regional ranking implied by (CP8) is perhaps a little surprising, inasmuch as Manitoba-Saskatchewan rather than the Atlantic Provinces falls at the bottom of the earnings list. Holding the industrial mix constant, in (CP9), yields the ranking one would have predicted for the time (1970): British Columbia, Ontario, Alberta, Quebec, Manitoba-Saskatchewan, the Atlantic Provinces. That this pattern should persist in the face of considerable standardization says much about the profoundness of regional disparity in Canada. As for TYPE, the 11.5% earnings advantage of metropolitan-area residents in (CP9) appears generally consistent with expectations.

If geographic mobility, the supply of information, and competition for employment were both perfect and costless, one would expect the coefficients of GEO and TYPE to be insignificant. It may be, of course, that the observed geographic and metropolitan-versus-rural-and-small-town differentials are really of an equalizing nature--the competitive outcome of varying tangible and intangible benefits and costs. Equation (CP9) then implies that the Atlantic Provinces supply the largest, and

British Columbia the smallest, real amenity total. It would surely be presumptuous to attempt an objective assessment of this proposition. One may say, comparing (CP8) and (CP9), that the net effect of equalizing differentials and market imperfection is to lower the estimated return to schooling by 0.75 percentage points. Together, GEO and TYPE explain an additional 2.4% of the earnings variance, or about one-third of the amount ascribed to schooling in (CS1).

IND adds a further 4.3% to the value of R^2.[40] Seven of its nine coefficients are significant. Hence, GEO, TYPE, and IND, at a minimum, contribute almost as much (6.7%) as S at its maximum (7.3%). When S is dropped from (CP9),[41] R^2 falls by 5.5 percentage points, indicating the minimum effect of the variable.[42] Of course, schooling does not pretend to measure the individual's total stock of human capital. If the latter is given by S, P, and PSQ, we may estimate its contribution from (CP1) at 21.3%. The market-imperfection variables have about one-third the explanatory power. They lower the implied rate of return to schooling by almost 2 percentage points.

The negative coefficient obtained for MAJ suggests that, on average, individuals pay a premium for being self-employed. The size of the premium may actually be somewhat larger than is indicated here, since one would expect the present coefficient to be biased upwards through the inclusion in earnings of some returns to non-human capital. On the other hand, because the self-employed category is extremely heterogeneous, the average figure may not be especially useful.

Family-status variables. The results of adding HEAD, FAMSIZ, and USMAR are displayed in Equation (CP10), Table 9. These variables are included here primarily for descriptive purposes, since we have not surveyed any rigourous theoretical arguments for their insertion. One might speculate that family and marital responsibilities could have some effect upon the individual's "reservation wage" during periods of job search. Those who have held out for a high wage at some time in the past, either because of perceived high subsistence requirements or because of available support from secondary earners, will tend to record high current incomes as a result. Discrimination in favour of married family heads may also be a factor. One should nevertheless be on guard against the strong likelihood that the variables in question are endogenous. Earnings may very well predetermine family status. At the very least, earnings and family status may be related solely through a common dependence upon some unmeasured quality of the individual.

At any rate, HEAD is uniformly significant with a large coefficient. USMAR is significant at the 0.05 level or better in all but Equation (CP9). FAMSIZ is nowhere significant in Table 9, but it becomes so in (CP14) and (CP15), Table 10, where WTIME has been deleted. Hours of work apparently interact with size of family to create a link between the latter variable and earnings, though size of family bears no relationship to the implicit wage. HEAD, USMAR, and FAMSIZ together account for a modest 1.5% of total earnings variance.

TABLE 9

REGRESSION ESTIMATES[a] OF THE EXPANDED EARNINGS FUNCTION, II

Ind. Variable	Equations[b] (dependent variable = INC)					
	(CP10)[c]		(CP11)[c]		(CP12)[c]	
Constant	.6007	(23.9)	.6422	(24.8)	.6421	(23.9)
S	.0675	(45.9)	.0653	(43.2)	.0651	(42.2)
P	.0416	(34.2)	.0414	(33.8)	.0415	(33.9)
PSQ	-.0007	(30.9)	-.0007	(30.7)	-.0007	(30.8)
WTIME	.6498	(81.1)	.6494	(81.1)	.6488	(80.9)
LENC2	-.0021	(0.79)	-.0191	(0.71)	-.0188	(0.70)
LENC3	.0075	(0.53)	.0086	(0.61)	.0085	(0.60)
LENC4	.0811	(4.56)	.0843	(4.73)	.0833	(4.66)
HEAD	.2404	(7.69)	.2360	(7.55)	.2309	(7.39)
FAMSIZ	-.0020	(0.81)	-.0019	(0.79)	-.0018	(0.75)
USMAR	.0590	(1.93)	.0683	(2.12)	.0696	(2.28)
IM1	-	-	.0268	(1.21)	.0267	(1.20)
IM2	-	-	-.0245	(1.84)	-.0190	(1.35)
IM3	-	-	-.1050	(4.70)	-.0856	(3.67)
LAN2	-	-	-.1091	(5.45)	-.1195	(5.43)
LAN3	-	-	.0150	(0.72)	-.0004	(0.26)
LAN4	-	-	-.0282	(0.65)	-.0321	(0.73)
ETH2	-	-	-	-	.0054	(0.44)
ETH3	-	-	-	-	-.0202	(1.02)
ETH4	-	-	-	-	-.1059	(2.08)
ETH5	-	-	-	-	.2301	(5.65)
ETH6	-	-	-	-	-.0460	(0.84)
ETH7	-	-	-	-	-.0512	(2.22)
REL2	-	-	-	-	.0141	(1.13)
REL3	-	-	-	-	-.0724	(3.12)
REL4	-	-	-	-	.0166	(0.88)
R^2	.467		.469		.470	

[a]Main sample, 22,682 observations

[b]The first figure in each set is a regression coefficient; the second, in parentheses, is the corresponding t ratio, written in absolute terms.

[c]Included but not shown are GEO, TYPE, IND, and MAJ.

Personal-background variables. Although the four characteristics identified here--that is, IM, LAN, ETH, and REL--appear to contribute negligibly to earnings inequality "at the margin," individual coefficients supply a fair amount of useful information. As might be expected, recent immigrants (IM3) suffer a modest earnings disadvantage (8.6% vis a vis the reference group in (CP12)), but those who have lived in the country for some time do approximately as well as the Canadian born. Unilingual francophones (LAN2) earn 11-12% less than unilingual anglophones and less, even, than individuals who have no fluency in either English or French (LAN4). At the same time, biligualism (LAN3) does not seem to confer any significant advantage. Adherence to a non-Christian religious faith (REL3) signals below-average earnings.

Of the six coefficients for ethnic group, three are significant at the 0.05 level or better. Given the standardization enforced in (CP12), we find that Jews in the sample (ETH5) earn an average of 23.0% more than the reference group, Chinese and Japanese (ETH4), 10.6% less, and Negro, West Indian and "other" (ETH7), 5.1% less. Native Indians (ETH6) also suffer a disadvantage, but it is not statistically significant.

One should not assume, however, that the preceding ethnic coefficients measure the full extent of any discrimination which may be present. There is, first of all, some degree of multicollinearity between ETH and each of the other three background variables IM, LAN, and REL. Secondly, it must be remembered that in (CP12), as in most of the other earnings functions, time worked is held constant. Discrimination may well manifest itself more significantly through hiring, turnover, and so on than through the payment of differentiated wages.

Table 10 therefore presents some further evidence. We see in (CP13) that removing IM, LAN, and REL does not have much overall effect, but it does lower the coefficient of ETH5 rather markedly. The reason is simple: as shown in Table 24 (Appendix IIIA), ETH5 and REL3 are practically the same variable, since most non-Christians in the sample are ethnically Jewish. In fact, the coefficient of ETH5 in (CP13) is virtually the algebraic sum formed by the coefficients of ETH5 and REL3 in (CP12). Removing WTIME has a profound effect on the coefficient of ETH6. The disadvantage borne by Native Indians does indeed appear to stem much more from employment than from wage rates. On average, native people earn 34-35% less than those in the reference group. Overall in (CP15), four of the six ethnic coefficients turn out to be significant.

Variable returns to schooling. By including only the linear term S in Table 8-10, we have so far dictated a constant rate of return to schooling. Table 11 relaxes this assumption by re-introducing the squared term SSQ and the experience interaction XSP. As before, the coefficient of SSQ is both positive and highly significant, implying that the rate of return increases with the level of schooling. The coefficient of S is driven to insignificance. That of XSP remains significantly negative. Thus even after extensive standardization,

TABLE 10

REGRESSION ESTIMATES OF THE EXPANDED EARNINGS FUNCTION, III

Ind. Variable	Equations[b] (dependent variable = INC)					
	(CP13)[c]		(CP14)[c]		(CP15)[c]	
Constant	.6177	(23.7)	.2554)	(8.51)	.2243	(7.70)
S	.0667	(44.5)	.0732	(41.9)	.0749	(44.2)
P	.0418	(34.4)	.0608	(44.5)	.0614	(45.4)
PSQ	-.0007	(31.2)	.0011	(43.4)	.0011	(44.3)
WTIME	.6492	(80.9)	-	-	-	-
LENC2	-.0213	(0.79)	-.0418	(1.37)	-.0451	(1.48)
LENC3	.0076	(0.54)	.0093	(0.58)	.0090	(0.56)
LENC4	.0800	(4.50)	.1020	(5.02)	.1010	(5.01)
HEAD	.2334	(7.47)	.3134	(8.84)	.3164	(8.91)
FAMSIZ	-.0021	(0.86)	-.0113	(4.09)	-.0116	(4.22)
USMAR	.0655	(2.14)	.1375	(3.96)	.1341	(3.85)
IM1	-	-	-.0261	(1.04)	-	-
IM2	-	-	-.0003	(0.02)	-	-
IM3	-	-	-.1056	(3.99)	-	-
LAN2	-	-	-.1402	(5.62)	-	-
LAN3	-	-	-.0157	(0.84)	-	-
LAN4	-	-	-.0204	(0.41)	-	-
ETH2	-.0042	(0.39)	-.0102	(0.72)	-.0059	(0.48)
ETH3	-.0158	(0.86)	-.0299	(1.33)	-.0284	(1.35)
ETH4	-.1362	(2.72)	-.1075	(1.86)	-.1396	(2.46)
ETH5	.1649	(4.73)	.2424	(5.25)	.1778	(4.50)
ETH6	-.0362	(0.66)	-.3510	(5.67)	-.3440	(5.56)
ETH7	-.0771	(3.65)	-.0497	(1.90)	-.0785	(3.28)
REL2	-	-	.0030	(0.21)	-	-
REL3	-	-	-.0785	(2.98)	-	-
REL4	-	-	.0038	(0.18)	-	-
R^2	.468		.317		.315	

[a]Main sample, 22,682 observations

[b]The first figure in each set is a regression coefficient; the second, in parentheses, is the corresponding t ratio, written in absolute terms

[c]Also included but not shown are GEO, TYPE, IND, and MAJ

TABLE 11

THE EXPANDED EARNINGS FUNCTION WITH A VARIABLE RATE OF RETURN (EQUATION (CP16)) [a]

Ind. Variable	Coefficient	(t ratio)[b]	Ind. Variable	Coefficient	(t ratio)[b]
Constant	.8626	(14.5)	IND9	-.0051	(0.22)
			IND10	-.1581	(11.6)
S	.0057	(0.74)			
SSQ	.0033	(12.0)			
			MAJ	-.0677	(4.14)
P	.0498	(24.2)			
PSQ	-.0008	(30.7)	HEAD	.2151	(6.92)
			FAMSIZ	-.0017	(0.69)
			USMAR	.0703	(2.31)
XSP	-.0005	(4.65)			
WTIME	.6482	(81.2)	IM1	.0259	(1.17)
			IM2	-.0312	(2.22)
			IM3	-.1036	(4.46)
LEN2	.0012	(0.44)			
LEN3	.0352	(2.48)			
LEN4	.1049	(5.89)	LAN2	-.1455	(6.63)
			LAN3	-.0182	(1.11)
			LAN4	-.0979	(2.23)
GEO1	-.2342	(13.4)			
GEO2	-.0246	(1.52)			
GEO4	-.1458	(8.63)	ETH2	.0052	(0.42)
GEO5	-.0499	(2.88)	ETH3	-.0367	(1.86)
GEO6	.0434	(2.85)	ETH4	-.1270	(2.51)
			ETH5	.2081	(5.14)
			ETH6	-.0967	(1.77)
TYPE	.1132	(11.6)	ETH7	-.0689	(3.00)
IND1	-.6731	(30.3)			
IND2	.0300	(0.86)	REL2	.0124	(1.00)
IND3	-.4295	(7.23)	REL3	-.0782	(3.39)
IND4	.1836	(6.74)	REL4	.0009	(0.05)
IND6	.0474	(3.05)			
IND7	.0305	(2.10)			
IND8	-.1533	(11.5)	R^2	.476	

[a]Estimated for the main sample, 22,682 observations.

[b]Absolute values.

experience profiles continue to exhibit convergence. At mean levels of schooling and experience, the estimated return to schooling (dINC/dS) is 6.0%.

Inserting SSQ and XSP in (CP16), Table 11, raises the R^2 by 0.6 of a percentage point. One might therefore be tempted to conclude that variation in the rate of return to schooling is not a very important source of earnings inequality. One cannot assume, however, that all variation in the rate of return expresses itself through SSQ and XSP. Much may be left in the residual. Although Mincer develops a way of partitioning the residual variance to obtain a maximum estimate of the component associated with variable returns,[43] his argument is inapplicable here because it assumes the independence of S and r^e. We have no recourse, it seems, but to account explicitly for variation in the rate of return through the use of additional determinants. The interactions model reported below pursues this problem.

Otherwise, the re-introduction of SSQ and XSP vaults three more variables into the "significant" category, namely: LENC3 (6 months - 3 years vocational training), IM2 (immigrated 1946 - 1965), and LAN4 (neither English nor French). GEO2 (Quebec residence) becomes insignificant. Comparing (CP15) and (CP16), one can see that the general pattern of coefficients is not much affected.

The occupational dimension. It has been argued[44] that including occupation in the earnings function along with schooling will necessarily bias downward the estimated rate of return, since individuals appear to reap the benefit of their schooling investment by moving upward through the occupational hierarchy. Holding occupation constant thus imposes an unnatural constraint. Nevertheless, it seems useful to examine the occupational dimension of earnings, not only for descriptive purposes, but also in order to test the empirical significance of the preceding objection.

Its practical validity must depend to a great extent on how "occupation" is defined. As usual, the researcher is very much at the mercy of the data. If the available categorization scheme rests on hierarchical factors such as the level of training, the degree of status, or the span of responsibility, then the bias problem just mentioned will be more severe than if the system is grounded in some abstract analysis of work function, the nature of the industry, or the type of good or service produced. In the latter case, occupational wage differentials are again likely to be of the equalizing variety, or else they are the result of noncompetitive forces.

The particular categorization scheme embodied in the PUS data is not easy to characterize in the preceding terms. Status, function, and industry all seem to play a role.[45] The headings are broad (since there are only twelve used here), and all would appear to admit individuals with widely varying levels of schooling. Schooling and occupation, as currently defined, are nonetheless correlated to a degree. It seems prudent therefore merely to let the results speak for themselves. The effects of adding occupation to the earnings function are displayed in Table 12.

TABLE 12

THE EFFECTS OF OCCUPATION[a]

Ind. Variable	Equations[b] (dependent variable = INC)							
	(CP17)		(CP18)		(CP19)[c]		(CP20)	
CONSTANT	.4382	(15.7)	.6850	(12.8)	.7868	(10.4)	.2681	(4.42)
S	.0534	(32.5)	.0255	(4.57)	.0236	(2.48)	.0281	(4.42)
SSQ	-	-	-	-	.0006	(1.81)	-	-
P	.0531	(45.8)	.0529	(45.7)	.0474	(23.3)	.0783	(61.6)
PSQ	-.0009	(40.4)	-.0009	(40.5)	-.0007	(29.7)	-.0013	(56.8)
XSP	-	-	-		-.0005	(4.42)	-	-
WTIME	.6596	(81.8)	.6587	(81.9)	.6420	(81.6)	-	-
OC1	.6743	(25.9)	.5229	(5.61)	.4608	(4.84)	.5869	(5.53)
OC2	.5103	(18.3)	-.0305	(0.30)	-.1042	(1.02)	-.0492	(0.43)
OC3	.4628	(14.1)	-.3414	(2.19)	-.1727	(1.11)	-.5693	(3.20)
OC4	.5151	(13.7)	-.8489	(6.66)	-.7206	(5.73)	-.8714	(6.00)
OC5	.2340	(10.1)	.0814	(0.94)	.0468	(0.56)	.0646	(0.66)
OC6	.2298	(10.7)	-.1384	(1.90)	-.0031	(0.04)	-.0178	(0.21)
OC8	-.2842	(13.2)	-.4812	(7.10)	-.1111	(1.60)	-.5195	(6.73)
OC9	.2618	(13.7)	.0604	(0.99)	.0525	(0.88)	.0784	(1.12)
OC10	.3303	(15.6)	.1173	(1.76)	.0698	(1.06)	.0656	(0.86)
OC11	.2117	(9.03)	-.0285	(0.35)	-.0318	(0.41)	-.0634	(0.67)
OC12	.2456	(11.4)	.1783	(2.69)	.0919	(1.42)	.1096	(1.45)
XSOC1	-	-	.0210	(2.63)	.0154	(1.88)	.0237	(2.61)
XSOC2	-	-	.0490	(6.00)	.0431	(5.09)	.0566	(6.08)
XSOC3	-	-	.0635	(5.88)	.0495	(4.50)	.0819	(6.68)
XSOC4	-	-	.1049	(11.0)	.0962	(9.98)	.1140	(10.5)
XSOC5	-	-	.0196	(2.33)	.0132	(1.62)	.0269	(2.81)
XSOC6	-	-	.0397	(5.45)	.0242	(3.41)	.0392	(4.72)
XSOC8	-	-	.0226	(2.97)	.0197	(2.71)	.0318	(3.67)
XSOC9	-	-	.0234	(3.53)	.0135	(2.13)	.0289	(3.82)
XSOC10	-	-	.0246	(3.38)	.0144	(2.08)	.0388	(4.68)
XSOC11	-	-	.0277	(3.13)	.0156	(1.84)	.0364	(3.61)
XSOC12	-	-	.0098	(1.42)	.0098	(1.45)	.0181	(2.28)
R^2	.436		.440		.495		.274	

[a]Estimated for the main sample, 22,682 observations.

[b]The first figure in each set is a regression coefficient; the second, in parenthesis, is the corresponding t ratio, written in absolute terms.

[c]Also included but not shown are LENG, GEO, TYPE, IND, MAJ, HEAD, FAMSIZ, USMAR, IM, LAN, ETH, and REL.

The eleven intercept dummies in (CP17) raise the level of R^2 by 5.4 percentage points, compared with (CP5), and lower the implied rate of return to schooling from 7.8% to 5.3%. The latter change represents the maximum extent of the possible bias. If it were the true extent, it could also be interpreted as measuring that component of the return to schooling which must be realized through occupational mobility. Doubtlessly, however, there exists some return to occupational mobility which is merely correlated with but not dependent upon the level of schooling. As one might easily have forecast, managerial personnel (OC1) rank at the top of the earnings scale, followed by workers in health care (OC4). Farm and other primary workers (OC8) rank lowest, preceded by service workers (the reference group, OC7).

Equations (CP18)-(CP20) add the vector of interaction terms XSOC. (CP19) includes the collection of variables treated earlier in Table 11; (CP20) is identical to (CP18) except for the deletion of WTIME. By adding the respective coefficients of XSOC to the coefficient of S, one may compute the set of intra-occupational rates of return. These are not, of course, the rates of return that individuals secure, having chosen to *enter* a particular occupation. They measure instead the rewards to educational upgrading *within* a particular category. Hence the large figure implied for workers in health care (XSOC4: 0.0255 + 0.1049 = 0.1304) must simply reflect unusual steepness in the earnings gradient across schooling levels in this field. Teaching (XSOC3) stands out in a similar fashion.

Occupation does appear to capture some variation in the rate of return, for in (CP19) the coefficient of SSQ becomes insigifnicant. Although the interaction terms add very little to the R^2, they are jointly significant in an F test at the 0.01 level. Permitting hours of work to vary, in (CP20), does not change the general pattern of these coefficients; but it does increase their values, as the employment factor becomes incorporated in the estimated rates of return. Most of the intercept coefficients fall algebraically, since the earnings-schooling gradients pivot upward to accommodate the rearranged scatter of observations.

Summary. Now that we have looked in detail at all the variable groups considered for inclusion in the earnings function, it is necessary to conduct a broad comparison of their quantitative influence. For this purpose Table 13 presents a decomposition of the explained earnings variance (inequality) and a set of F statistics pertaining to the variable groups. These F statistics are more useful in the current context than the standard t ratios given earlier, since the latter, being in part dependent upon the choice of a reference group, are bound to be somewhat arbitrary. As noted previously, we cannot avoid a certain degree of arbitrariness involving the order in which variables enter the regression equations. Since the order shown in Table 13 tends to favour (gives the "benefit of the doubt" to) the orthodox human-capital variables by introducing them first, we must pay some attention,

TABLE 13

THE EXPLANATORY POWER AND SIGNIFICANCE OF VARIABLES IN THE EXPANDED EARNINGS FUNCTIONS

Variable Group	Variance Increment[a]			F Statistic	
	Upon Addition	Percent of Exp. Var.[b]	Upon Deletion	Upon Addition	Upon Deletion
S	.07332	14.82	.00014	11341.00	6.40
P PSQ	.14011	28.33	.01999	17443.69	457.06
WTIME	.16873	34.12	.14872	10012.27	6800.76
LENC	.00248	0.50	.00089	48.73	13.57
GEO TYPE	.02476	5.01	.01093	228.49	83.30
IND MAJ	.04210	8.51	.02213	234.76	101.20
HEAD FAMSIZ USMAR	.01558	3.15	.01198	251.91	182.61
IM	.00064	0.13	.00042	10.33	6.40
LAN	.00122	0.25	.00128	19.64	19.51
ETH	.00081	0.16	.00066	6.51	5.03
REL	.00033	0.07	.00031	5.49	5.30
SSQ XSP	.00632	1.28	.00075	150.16	17.15
OC XSOC	.01817	3.67	.01817	37.77	37.77
Total	.49457	100.00	-	-	-

[a]Change in R^2. Variable groups were added to the regression in the order shown and then deleted singly.

[b]Change in R^2 upon addition, divided by maximum R^2 with all variables included (x 100)

as we did earlier, to the alternatives. The table thus reports the change in R^2 observed upon the deletion of each variable or variable group from the full model.

It is clear from Table 13, if not from all the previous results, that WTIME is by far the most important explanatory variable. The decision to explore this variable further in Chapters IV and V thus appears well founded. Experience (or more agnostically, the "life-cycle factor") was included early and is important upon addition but very much less so upon deletion. The linear term for schooling behaves similarly. One should note, however, that the presence of SSQ, XSP, and XSOC in the full model predisposes this result. When all the human-capital variables and their interactions are deleted, the R^2 falls by 0.042; the F statistic for their joint significance is 99.69. Conversely, when the "unorthodox" variables GEO through OC are deleted, the R^2 falls by 0.105; and the corresponding F statistic is 83.16.

Broadly speaking, geographic and industrial factors seem to play an important role in earnings and inequality determination--very nearly as important, perhaps, as that of schooling. Family status is associated with earnings, although one cannot be confident about the direction of causality. The personal-background variables identified here account for a very small proportion of total inequality, at least insofar as wage rates are concerned. Nevertheless, the significance of individual coefficients shows that some small groups may have strongly divergent earnings experiences.

TABLE 14

RATES OF RETURN TO SCHOOLING IMPLIED BY VARIOUS SPECIFICATIONS OF THE EARNINGS FUNCTION

Equation Number	Estimated Return (%)	Details of Specification [a]
(CS1)	6.95	Includes S only
(CP1)	8.91	Adds P, PSQ
(CP2)	8.66[b]	Adds SSQ, XSP
(CP5)	7.75	Includes WTIME; excludes SSQ, XSP
(CP17)	5.34	Includes OC
(CP7)	7.63	Adds LENC, excludes OC
(CP8)	6.88	Adds GEO, TYPE
(CP9)	7.05	Adds IND, MAJ
(CP10)	6.75	Adds HEAD, FAMSIZ, USMAR
(CP12)	6.51	Adds IM, LAN, ETH, REL
(CP14)	7.32	Excludes WTIME
(CP16)	6.03[b]	Re-inserts WTIME, SSQ, XSP

[a]Changes noted are cumulative

[b]Calculated at mean levels of schooling and experience

As a final matter, it seems useful to compare, all at once, the schooling coefficients obtained from various specifications of the earnings function. These are collected in Table 14. The largest implied rates of return occur with hours of work free to vary; the smallest, when occupation is held constant. With hours fixed, the range is from 6.03% to 7.75%; with hours variable, it is from 6.95% to 8.91%. In neither case does the degree of uncertainty seem especially serious from a policy point of view.

If one were to add a correction for economic growth[46]--say, 2.5%--as does the previously cited Statistics Canada study,[47] the preceding figures would increase accordingly. In the comparable (time-variable) case, they tend to exceed the Statistics Canada estimate of approximately 8%. However, the latter takes into account the direct private and social costs of education, which are ignored by the current procedures. The present estimates imply returns lower than found by Podoluk for Canada a decade earlier and lower than reported by Mincer for the United States.

An Interactions Model

At several points in preceding chapters we have considered the interactions specification put forward by Haessel and Kuch. It will be recalled that these authors attempt to explain possible disparities in the rate of return to human capital by making them a function of certain independent variables. Since earnings are assumed to equal (at least in part) the product of human capital and its rate of return, the result, upon substitution for the latter, is an estimating equation displaying a number of interaction terms.

In selecting variables to explain the rate of return, Haessel and Kuch emphasize personal background and occupation.

Using the former, they investigate the problem of discrimination. The present study is more concerned, however, with the sort of market imperfection which may be captured by the variables "industry" and "place of residence." Hence, the following regression model is postulated:

$$INC_i = c_0 + r_i \cdot H_i + b_4 \cdot WTIME_i + u_i$$

$$r_i = a_0 + a_1' \cdot GEO_i + a_2' \cdot IND_i$$

$$H_i = h_0 + h_1 S_i + h_2 P_i + h_3 PSQ_i \quad ,$$

where a_1' and a_2' are row vectors of coefficients multiplying the column vectors IND_i and GEO_i, which describe individual i. As in previous notation, r_i stands for the average rate of return on units of human capital, the total accumulation of which is given by H_i; and u_i is an error term with classical properties. The remaining lower-case symbols are scalar coefficients. Upon substitution into the first equation we obtain:

$$\begin{aligned} INC_i = {} & (c_0 + a_0 h_0) + a_0 h_1 S_i + a_o h_2 P_i + a_0 h_3 PSQ_i + B_4 WTIME_i \\ & + a_1' h_0 GEO_i + a_2' h_0 IND_i + a_1' j_1 XSGEO_i + a_2' h_1 XSIND_i \\ & + a_1' h_2 XPGEO_i + a_1' h_3 XPSQGEO_i + a_2' h_2 XPIND_i \\ & + a_2' h_3 XPSQIND_i + u_i \quad , \end{aligned}$$

where the interaction terms are as defined in Table 3. The regression coefficients may be defined implicitly by writing

$$\begin{aligned} INC_i = {} & b_0 + b_1S_i + b_2P_i + b_3PSQ_i + b_4WTIME_i \\ & b_5'GEO_i + b_6'IND_i + b_7'XSGEO_i + b_8'XSIND_i \\ & b_9'XPGEO_i + b_{10}'XPSQGEO_i + b_{12}'XPIND \\ & + b_{13}'XPSQIND_i + u_i \quad . \end{aligned}$$

Here, b_0 through b_4 are scalars; b_5' through b_{13}' are row vectors.

The preceding equation is amenable to ordinary least squares estimation by virtue of the fact that the expressions for r_i and H_i are assumed nonstochastic. Haessel and Kuch show that if random components other than u are present, the model will be subject to heteroskedasticity. They consequently develop an asymptotically efficient (maximum-likelihood) estimation procedure.[48] Owing to the computational burden involved in treating the present sample, this refinement is not pursued here. We must therefore be somewhat cautious in accepting the derived standard errors, although the estimated coefficients are presumably unbiased.

From the coefficients it is possible to obtain estimates of the return to schooling within a given region or industry. One need only compute

$$d\ INC_i / dS_i \ = \ b_1 + b'_7(dXSGEO_i / dS_i) + b'_8(dXSIND_i / dS_i) \quad .$$

Note, however, that this rate of return is not quite the same thing as r_i, the analytical device used above. The latter is the rate of return to a unit of human capital; the former is the rate of return to a (time) unit of schooling.

Results appear in Table 15. The schooling interactions shown in (C11) contribute only 0.004 to the value of R^2, though as a group they are highly significant.[49] The vectors XSGEO and XSIND, taken in that order, are significant individually as well. Over regions, as shown by the former, the implied rate of return varies from 7.5% in Atlantic Canada to 4.3% in British Columbia (for workers in the reference industry, manufacturing). Since these regions are generally regarded as being at or near opposite ends of the scale with respect to levels of education and human-capital scarcity, this outcome seems consistent with ordinary demand-and-supply inferences. Over industries, the range is a little larger than over regions--about 4.7 percentage points. As in the case of occupation, however, it may be deemed somewhat improper to hold industry constant in estimating returns. The relevant opportunity wage need not be found in the industry within which the individual is currently employed.

This objection is perhaps less serious with respect to the experience interactions. Because workers tend to give up mobility

TABLE 15

THE INTERACTION OF SCHOOLING AND EXPERIENCE WITH INDUSTRY AND PLACE OF RESIDENCE [a]

Ind. Variable	Equations[b] (dependent variable = INC)			
	(CI1)		(CI2)	
Constant	.8142	(22.1)	.7399	(15.1)
S	.0571	(18.2)	.0593	(18.0)
P	.0519	(45.5)	.0559	(22.1)
PSQ	-.0008	(38.3)	-.0009	(18.4)
WTIME	.6807	(85.0)	.6807	(84.9)
GEO1	-.3860	(7.65)	-.2770	(3.50)
GEO2	-.1266	(3.87)	-.2433	(4.67)
GEO4	-.1568	(2.87)	.0444	(0.52)
GEO5	-.0531	(0.89	-.1005	(1.13)
GEO6	.1804	(3.39)	.3375	(4.32)
IND1	-.7041	(10.8)	-.3311	(3.10)
IND2	-.0832	(0.79)	-.0779	(0.48)
IND3	-.6889	(4.04)	-.8619	(3.02)
IND4	.1581	(1.75)	.1828	(1.29)
IND6	.1626	(3.11)	.2040	(2.44)
IND7	.0312	(0.62)	.1266	(1.60)
IND8	-.1965	(4.05)	-.2183	(3.17)
IND9	-.0731	(0.77)	-.0199	(0.15)
IND10	-.5111	(12.7)	-.3602	(5.86)
XSGEO1	.0179	(3.60)	.0180	(3.36)
XSGEO2	.0059	(1.88)	.0099	(2.94)
XSGEO4	.0008	(0.16)	-.0054	(0.94)
SCGEO5	.0008	(0.15)	.0043	(0.73)
XSGEO6	-.0134	(2.82)	-.0182	(3.61)
XSIND1	-.0054	(0.76)	-.0157	(2.03)
XSIND2	.0091	(0.76)	.0040	(0.32)
XSIND3	.0346	(1.65)	.0486	(2.12)
XSIND4	.0038	(0.42)	-.0006	(0.07)
XSIND6	-.0129	(2.35)	-.0173	(2.89)
XSIND7	.0006	(0.12)	-.0022	(0.41)

Ind. Variable	Equations[b] (dependent variable = INC)			
	(CI1)		(CI2)	
XSIND8	.0038	(0.81)	.0029	(0.58)
XSIND9	.0087	(1.10)	.0113	(1.32)
XSIND10	.0337	(9.49)	.0298	(7.77)
XPGEO1	-	-	-.0151	(3.46)
XPGEO2	-	-	.0052	(1.86)
XPGEO4	-	-	-.0111	(2.78)
XPGEO5	-	-	-.0017	(0.40)
XPGEO6	-	-	-.0063	(1.63)
XPIND1	-	-	-.0193	(3.98)
XPIND2	-	-	.0114	(1.19)
XPIND3	-	-	-.0001	(0.00)
XPIND4	-	-	.0074	(0.91)
XPIND6	-	-	.0050	(1.18)
XPIND7	-	-	-.0052	(1.23)
XPIND8	-	-	.0050	(1.49)
XPIND9	-	-	-.0119	(2.17)
XPIND10	-	-	-.0093	(2.80)
XPSQGEO1	-	-	.0003	(3.93)
XPSQGEO2	-	-	.0001	(1.02)
XPSQGEO4	-	-	.0002	(2.33)
XPSQGEO5	-	-	.0001	(0.92)
XPSQGEO6	-	-	.0001	(0.79)
XPSQIND1	-	-	.0003	(3.06)
XPSQIND2	-	-	-.0003	(1.68)
XPSQIND3	-	-	.0001	(0.37)
XPSQIND4	-	-	-.0002	(1.38)
XPSQIND6	-	-	-.0002	(2.12)
XPSQIND7	-	-	.0001	(0.85)
XPSQIND8	-	-	-.0001	(1.84)
XPSQIND9	-	-	.0002	(2.44)
XPSQIND10	-	-	.0001	(2.16)
R^2	.455		.458	

[a]Estimated for the main sample, 22,682 observations

[b]The first figure in each set is a regression coefficient; the second, in parentheses, is the corresponding t ratio, written in absolute terms.

as they gain experience, rates of return to the latter form of human capital within particular regions and industries may be of definite practical relevance. Like the schooling interactions in (CI1), those involving experience in (CI2) add very little to the R^2, but enough to be judged significant in an F test at the 0.01 level.[50] The "return" to an additional year of experience is lowest at the (national) mean in Manitoba-Saskatchewan (1.24%) and highest in Alberta (1.72%). It is lowest in agriculture (0.88%) and highest in fishing (1.88%).[51]

Although rates of return to schooling and experience do appear to vary across regions and industries, it cannot be claimed that such variation contributes very strongly to the prevailing level of earnings inequality. Whereas, region and industry are important in themselves,[52] they do not have much effect on the earnings potency of discretionary human-capital investment. If such variation in the rate of return is indeed an important source of inequality, better data, with groups more narrowly defined than at present, will obviously be needed to establish the fact.

APPENDIX IIIA

THE WORKING SAMPLE: DISTRIBUTIONS OF SELECTED CHARACTERISTICS

TABLE 16

INDIVIDUAL INCOMES BY SIZE CATEGORY

Size Category ($'s)	Numbers of Individuals	
	Employment Income	Total Income
0- 999	1136	748
1,000- 1,999	1254	1108
2,000- 2,999	1462	1343
3,000- 3,999	1740	1668
4,000- 4,999	2004	2010
5,000- 5,999	2368	2362
6,000- 6,999	2519	2465
7,000- 7,999	2440	2505
8,000- 9,999	3344	3495
10,000-11,999	1838	2076
12,000-14,999	1254	1380
15,000-17,999	536	606
18,000-24,999	428	493
25,000-34,999	199	220
50,000-74,999	42	44
75,000 or more	18	25
Total	22,682	22,682

TABLE 17

FAMILY INCOMES OF INDIVIDUALS BY SIZE CATEGORY

Size Category ($'s)	Number of Individuals
Loss	9
0	0
1- 999	123
1,000- 1,999	262
2,000- 2,999	529
3,000- 3,999	752
4,000- 4,999	1,020
5,000- 5,999	1,263
6,000- 6,999	1,512
7,000- 7,999	1,674
8,000- 9,999	3,374
10,000-11,999	2,889
12,000-14,999	2,891
15,000-19,999	2,080
20,000-24,999	745
25,000-34,999	410
35,000-49,999	194
50,000 or more	98
Nonfamily Individuals	2,857
Total	22,682

TABLE 18

SCHOOLING BY AGE GROUP

Level of Schooling[a]	Number of Individuals Aged					Total
	15-24	25-34	35-44	45-54	55+	
1	7	9	42	39	70	167
2	32	67	159	234	378	870
3	500	1261	1738	1647	1614	6760
4	939	1411	1219	983	752	5304
5	579	680	474	386	267	2386
6	922	923	599	472	311	3227
7	100	247	200	178	184	909
8	188	323	188	123	101	923
9	43	130	79	55	38	345
10	92	308	169	136	78	783
11	5	31	25	14	9	84
12	30	329	267	180	118	924
Total	3437	5719	5159	4447	3920	22,682

[a]1 = no schooling; 2 = grades 1-4; 3 = grades 5-8; 4 = grades 9-10; 5 = grade 11; 6 = grade 12; 7 = grade 13; 8 = 1-2 years university; 9 = 3-4 years university, without degree; 10 = 3-4 years university, with degree; 11 = 5 or more years university, without degree; 12 = 5 or more years university, with degree.

TABLE 19

SCHOOLING BY REGION

Schooling[a]	Numbers of Individuals					
	Atlantic	Quebec	Ontario	Manitoba-Sask.	Alberta	B.C.
1	36	26	61	15	15	14
2	113	346	163	74	31	43
3	625	2319	2297	638	394	487
4	443	1446	2033	456	407	519
5	232	704	753	229	206	262
6	157	537	1291	293	391	558
7	8	73	682	15	20	111
8	63	242	373	70	60	115
9	29	86	110	29	34	57
10	36	250	317	50	61	69
11	64	237	361	66	82	114
Total	1810	6302	8572	1937	1706	2355

[a]See footnote to Table 18.

TABLE 20

MEAN EARNINGS BY REGION AND LEVEL OF SCHOOLING

Level of Schooling[a]	Mean Earnings ($'s)						
	Canada	Atl.	Quebec	Ontario	Manitoba-Sask.	Alta.	B.C.
1	4090	2892	3690	5417	3076	3190	4177
2	4740	3367	4522	5709	3862	4932	5534
3	5889	4228	5696	6464	4924	6581	6931
4	6576	6576	6333	6927	5974	6497	7298
5	7227	6720	6753	7905	6443	7022	7840
6	7371	5757	7349	7785	5775	7282	7789
7	9157	7411	10403	9235	6992	7701	8537
8	8379	8130	8310	8633	9345	7107	7914
9	8356	5944	9153	8732	6915	11184	6698
10	11190	7982	10743	12501	9397	10422	10434
11	8470	8117	9425	8110	2635	5406	9541
12	16365	12015	14808	18804	14215	13524	17612
All Levels	7233	5472	6793	7963	6060	7306	8019

[a]See footnote to Table 18

TABLE 21

SCHOOLING BY INDUSTRY

Level of Schooling[a]	Numbers of Individuals Employed				
	Agricult.	Forestry	Fishing	Mining Petroleum	Manufac.
1	22	8	4	7	38
2	81	34	18	23	249
3	719	173	60	184	2007
4	320	81	29	180	1598
5	92	31	11	68	679
6	109	28	5	78	944
7	16	6	0	19	281
8	27	4	0	16	220
9	4	3	1	5	72
10	12	2	0	19	170
11	1	0	0	1	27
12	10	2	0	14	94
Total	1413	372	128	614	6379
	Constr.	Transp., Commun., Ut.	Trade	Finance	Services
1	22	11	17	1	37
2	120	113	92	12	128
3	961	824	944	87	801
4	616	744	962	123	651
5	226	356	493	106	324
6	286	442	653	209	473
7	56	86	157	119	169
8	66	106	154	97	233
9	19	43	65	22	111
10	16	56	76	53	379
11	5	8	7	6	29
12	12	49	29	39	675
Total	2405	2838	3649	874	4010

[a] See footnote to Table 18

TABLE 22

MEAN EARNINGS BY INDUSTRY AND LEVEL OF SCHOOLING

Level of Schooling[a]	Mean Earnings ($'s)				
	Agricul.	Forestry	Fishing	Mining Petroleum	Manufac.
1	2353	3907	875	6173	4140
2	2981	3672	2239	5097	5528
3	3984	5175	3557	7180	6337
4	4709	6001	3244	7860	6800
5	5052	7096	3786	8504	7172
6	4531	10970	4328	7713	7753
7	4403	8373	-	9226	9831
8	3905	9342	-	9585	8942
9	5057	2770	1000	12154	8919
10	4844	14320	-	13875	11813
11	1010	-	-	7900	8141
12	21463	8630	-	14150	13484
All Levels	4312	5931	3247	8038	7239
	Constr.	Transp., Commun., Ut.	Trade	Finance	Services
1	5700	4950	3779	5010	3970
2	5603	5352	4038	5678	3958
3	6454	6599	5880	6395	5056
4	6695	7440	6448	7839	5658
5	7231	8551	7148	8761	5980
6	7575	7917	7039	8569	6323
7	8599	10307	8271	10083	8277
8	6849	8324	8162	10414	8023
9	7795	9285	9221	7791	7498
10	12716	12201	13681	15202	9685
11	8892	10602	8750	9335	8144
12	12003	12580	12912	14656	17359
All Levels	6819	7656	6732	9300	8451

[a]See Footnote to Table 18

TABLE 23

OCCUPATION

	Occupational Category	Number of Individuals
1	Managerial and administrative	1241
2	Natural and social sciences, engineering	988
3	Teaching	635
4	Medicine and health care	407
5	Clerical	1702
6	Sales	2525
7	Service	1592
8	Farming and other primary	2236
9	Processing, fabricating, repairing	4868
10	Construction trades	2556
11	Transport equipment operation	1597
12	Arts, religion, other, and not stated	2335
Total		22,682

TABLE 24

ETHNIC AND RELIGIOUS GROUP

Ethnic Group	Religion				
	Protestant	Catholic and Orth.	Jewish and Other	No Religion	Total
1. British Is.	6949	1610	394	824	9777
2. W. European	1861	7302	240	329	9732
3. E. European	312	978	54	103	1437
4. Chinese and Japanese	52	18	32	66	168
5. Jewish	4	1	346	11	362
6. Nat. Indian	55	73	9	6	143
7. Other	191	685	114	63	1053
Total	9424	10667	1189	1402	22682

TABLE 25

PERIOD OF IMMIGRATION
TO CANADA

Period of Immigration	Number of Individuals
Before 1946	1025
1946 - 1965	3073
1966 or later	953
Canadian born	17631
Total	22682

APPENDIX IIIB

MISCELLANEOUS REGRESSIONS[1]

(CS2)	INC	=	.6077 (13.0)	+	.0688 S (25.0)	+	.0588 P (26.5)	-	.0010 PSQ (24.2)	+	.8776 WEEKS (45.9)
			- .2731 GEO1 (4.46)	-	.0569 GEO2 (1.38)	-	.2905 GEO4 (5.91)	-	.0669 GEO5 (1.35)	+	.0468 GEO6 (1.11)
			+ .0725 SPHG1 (1.25)	+	.0492 SPHG2 (1.09)	+	.0486 SPHG4 (1.02)	+	.0174 SPHG5 (0.30)	+	.0516 SPHG6 (0.92)
			+ .0192 SPHG7 (0.67)								
	R^2	=	.464		number of observations	=	5670				
(CS3)	INC	=	.5840 (12.4)	+	.0696 S (25.2)	+	.0599 P (26.9)	-	.0010 PSQ (24.6)	+	.8777 WEEKS (46.1)
			- .1446 SPHG1 (4.48)	-	.0011 SPHG2 (0.46)	-	.1425 SPHG4 (4.67)	-	.0311 SPHG5 (0.79)	+	.0971 SPHG6 (2.38)
			+ .0043 SPHG7 (0.16)								
	R^2	=	.458		number of observations	=	5670				
(CS4)	INC	=	.6237 (13.7)	+	.0683 S (25.0)	+	.0587 P (26.7)	-	.0010 PSQ (24.4)	+	.8775 WEEKS (46.2)
			- .2138 GEO1 - (6.49)	-	.02193 GEO2 (1.03)	-	.2562 GEO4 (8.03)	-	.0527 GEO5 (1.57)	+	.0755 GEO6 (2.56)
	R^2	=	.464		number of observations	=	5670				
(CS5)	INC	=	1.041 (34.8)	+	.0873 SCOST (26.7)	+	.0603 P (27.1)	-	.0010 PSQ (25.1)	+	.8827 WEEKS (46.5)
	R^2	=	.454		number of observations	=	5670				
(CS6)	INC	=	1.550 (93.7)	+	.0426 P (36.1)	-	.0008 PSQ (34.4)	+	.7143 WTIME (85.3)	+	.0044 LENC2 (0.15)
			+ .0481 LENC3 (3.20)	+	.1614 LENC4 (8.56)	-	.2576 GEO1 (14.0)	-	.1343 GEO2 (11.8)	-	.1586 GEO4 (8.98)
			- .0340 GEO5 (1.86)	-	.0661 GEO6 (4.11)	+	.1632 TYPE (16.2)	-	.7530 IND1 (34.5)	-	.0607 IND2 (1.64)
			- .5053 IND3 (8.16)	+	.2000 IND4 (6.86)	+	.0024 IND6 (0.15)	+	.0547 IND7 (3.53)	-	.1424 IND8 (10.0)
			+ .1556 IND9 (6.32)	+	.0202 IND10 (1.46)						
	R^2	=	.397		number of observations	=	22682				

[1]Figures in parentheses are t ratios, written in absolute terms.

NOTES

CHAPTER III

[1]For a complete description see Canada, Statistics Canada, Public Use Sample Tapes: User Documentation.

[2]One might think of using a "Tobit" procedure in this situation; however, such an approach will not be explored here. Zero earnings are not *per se* inconsistent with the model if k = 1. Yet, individuals are not generally observed to specialize in on-the-job training.

[3]There is, of course, the purely mechanical problem of expressing nonpositive earnings in logarithmic form. In any case, negative earnings are likely to be a transitory phenomenon for the individual, better ascribed to ownership of physical capital and to entrepreneurship than to human capital.

[4]This is not to say, unfortunately, that the sample consists *only* of workers in the private sector. Only those in "public administration and defence" (S.I.C. Division II) could be excluded.

[5]A coin flip in fact chose the second.

[6]The latter was $6574. See Canada, Statistics Canada, 1971 Census of Canada, Vol. III, pt. 6, Income of Individuals, Catalogue no. 94-768 (Ottawa: Statistics Canada, May 1975), p. 1, Table 19.

[7]*Schooling, Experience, and Earnings*, p. 90.

[8]Isolating these factors *completely* of course demands both slope and intercept dummies. Slope dummies are not provided here except in the form of one interaction between agriculture and years of schooling. In preliminary testing the insertion of this latter variable and the intercept dummy for agriculture lowered the schooling

by about 0.5 percentage points. This result implies that omitting farmers might cause an even larger divergence between the present findings and those of Mincer than is observed below.

[9]He reports: "The regression coefficients in the age cross-section were very close to those in the experience cross-section, but the multiple coefficients of determination were .02 - .03 lower in the age set. . . ." Ibid., p. 91, no. 7.

[10]Presumably, such individuals are no longer making positive gross investments. To represent their experience profiles may strictly require a nonsmooth function. The exponential form may be especially inappropriate since as we have seen, it never falls to zero.

[11]Ibid., p. 90.

[12]The Pareto distribution is given by

$$f(Y) = Ay^{-\alpha} \quad ,$$

where A and α are constants ($\alpha > 2$) and f(y) is the proportion of individuals with income greater than Y. If V represents the largest income in the population and U, the boundary of the open-ended class, the mean income in this interval is given by

$$\frac{\int_U^V AY^{-\alpha} \cdot YdY}{\int_U^V AY^{-\alpha}\, dY} = \frac{[A/(2-\alpha)]Y^{2-\alpha}\Big|_U^V}{[A/(1-\alpha)]Y^{1-\alpha}\Big|_U^V} = \frac{1-\alpha}{2-\alpha} U \quad ,$$

as long as V is large.

Fitting a Pareto curve to the distribution if INC within the sample yielded a value of 2.657 for α. This implies a mean of \$189,200.

[13]For example, if self-employment is like a lottery, with a few large gains and many small losses (relative to other opportunities), individuals who choose to enter may willingly pay a premium in the form of inferior returns. Those with a taste for self-direction may do the same.

[14]The open-ended class was dealt with in the manner explained above.

[15]The variables USMAR, HEAD, and FAMSIZ were used in making the required determination.

[16]Canada, Dominion Bureau of Statistics, Principal Taxes and Rates: Federal, Provincial and Selected Municipal Governments, 1970 (Ottawa: Queen's Printer, 1970).

[17]The necessary figures were obtained from Canada, Department of National Revenue, 1972 Taxation Statistics [1970 taxation year] (Ottawa Information Canada, 1972), p. 152, Table 16.

[18]The source was ibid., pp. 150-151, Table 15.

[19]A problem here is that GEO-CODE gives the individual's residence on July 1, 1971, not his residence for tax purposes in 1970. Some error may thus attach to recent interprovincial migrants.

[20]In the fifth class, 50 was used rather than 50.5.

[21]See, for example, Canada, Health and Welfare Canada, Characteristics of Low-Wage Earners in Canada, Social Security Research Report No. 01 (Ottawa: Information Service Department of National Health and Welfare, September, 1976); or United States, Bureau of the Census, Statistical Abstract of the United States, 1970 (Washington: U.S. Government Printing Office, 1970).

[22]It may happen, of course, that grades and years do not correspond, as students skip grades or fail to win promotion. Whereas, years of schooling measure investment costs, one may speculate that grades relate more closely the mastery of certain skills and, hence, to productivity. The adopted procedure thus leans, if at all, toward the latter interpretation.

[23]Canada, Statistics Canada, Data Processing Division, "Special Tabulations 12295A and 12295B" (unpublished, September, 1976). Place of highest grade was selected a priori instead of place of current residence because the former, being less distant in time and more intimately connected with the environmental factors determining education, seemed more likely to be a good predictor of schooling.

[24]This assumption and the one below match those of Haessel and Kuch, "Earnings in Canada."

[25]As one would expect, place of residence is correlated with the schooling predictor, place of highest grade. In the sample, correlation coefficients between corresponding elements of GEO and SPHG (see below in the text, or Table 3) average about 0.8.

[26]To be more precise, under the standard procedure S contains a measurement error which is likely to be correlated with the variables named. The analysis is similar to that presented in Appendix IIB.

[27]Mincer apparently uses age 14. See Schooling, Experience, and Earnings, p. 48, notes to Table 3.1.

[28]Mincer assumes age 5; others, age 6. This scaling affects not only the regression constant but also the coefficients adhering to the various nonlinear transformations of P.

[29]This is Mincer's procedure.

[30]For precise definitions see Canada, Statistics Canada, Dictionary of the 1971 Census Terms (Ottawa: Statistics Canada, 1972).

[31]Here and below, cf. Haessel and Kuch, "Earnings in Canada.

[32]These descriptions apply to the earnings function. Recall that the quadratic stems from a linear investment profile.

[33]This programme was written by Keith Wales formerly of the University of British Columbia Computer Centre.

[34]See Appendix IIA.

[35]Note that is the system used here to number regression equations, "C" stands for "Canada," and other alphanumeric characters for the estimation procedure or specification. Thus (CV4) corresponds to Mincer's (V4), and so on

[36]See Podoluk, Incomes of Canadians.

[37]Ibid.

[38]Since vocational training was not deducted in computing experience, it might be argued that some "double counting" of human capital takes place when LENC and P appear in the same regression. To avoid confusion, one must carefully interpret LENC as signifying only the intensity of investment in relation to the average subsumed under P.

[39]Haessel and Kuch "Earnings in Canada," use a dummy vector similar to SPHG, but they do not encounter the multicollinearity problem inasmuch as their sample consists entirely of individuals resident in Toronto or Montreal.

[40]The contribution of MAJ is negligible in comparison.

[41]See Appendix IIIB, Equation (C56).

[42]See, however, Table 14. "Minimum" relates only to the present subset of variables.

[43]Schooling, Experience, and Earnings, p. 56.

[44]See Chapter II.

[45]See Table 3, and for a detailed explanation, Canada, Statistics Canada, Occupational Classification Manual, Census of Canada, 1971 (Ottawa: Information Canada, 1972).

[46]See the discussion in Chapter II.

[47]Economic Returns to Education.

[48]The authors unfortunately do not report the extent to which their efficient estimates differ from those provided by ordinary least squares.

[49]$F = 15.36$.

[50] F = 5.63.

[51] One suspects that the varying payoff to experience may have something to do with the pace of technological change in the two industries. Experience counts least where change is rapid. Investigation of this hypothesis is nevertheless beyond the scope of the present study.

[52] Observe that, within the context of the interactive model, the intercept terms for region and industry explain differences in the rate of return on the individual's initial endowment of human capital.

CHAPTER IV

THE SIMULTANEOUS DETERMINATION OF HUMAN-CAPITAL INVESTMENT AND LABOUR SUPPLY

The investment models we have so far considered treat labour supply as an exogenous factor in earnings determination. The sole problem for the individual is to choose an investment profile which maximizes net discounted lifetime earnings, or "wealth." Since there is in effect only one good, wealth and utility maximization amount to the same thing. In pursuing this simple objective, the individual is further assumed to ignore all systematic variation in planned or in realized hours of work.[1] Hence, the work profile is not only exogenous but also constant over the life cycle.

Both assumptions appear untenable. Empirically, the work profile is somewhat peaked, rather than horizontal.[2] Though it would not be very difficult to incorporate this or any other exogenous shape into an amended wealth-maximization model, it remains to be shown whether the standard prediction of monotonically declining investment in human capital would continue to hold. Theoretically, it *is* difficult to ignore the repercussions of the labour-leisure choice. That choice presumably depends upon a utility function which includes time in the form of leisure as an argument. Yet time is also the lone or principal input in the production of human capital. The rational individual will no doubt wish to allocate his fixed endowment of time optimally among work, leisure, and investment. Decision-making will be simultaneous rather than sequential, contrary to our previous assumption. To understand such behaviour, we must apparently discard the firm-based notion of independence between consumption and investment[3] and extend the analysis from the maximization of lifetime earnings to the maximization of utility.[4]

At the same time, it is especially important to keep in mind a point raised earlier, in Chapter II--namely, that the rate of return to any form of human capital is not well defined unless some reference is made to hours of work. Moreover, if work and investment are planned simultaneously, rates of return are "doubly endogenous" in the sense that they depend not only upon total investment, as in the Becker model, but also upon the profile of hours. Though it is always possible to compute the rate of return to schooling ex post for a given cross section of individuals, such an estimate will not correspond, *even in equilibrium*, to the rate apprehended by these individuals if we assume the wrong hours profile.

The first section of this chapter surveys a small group of theoretical studies which explore the simultaneous determination of human-capital investment and labour supply. From the standpoint of later empirical application, it is chiefly important in reviewing this work to find the answers to a pair of broad questions. The first, already mentioned, is whether the endogeneity of individual labour supply might upset the proof that investment declines monotonically over the life cycle. If the optimal propensity to invest is ever rising, the human-capital interpretation of concave earnings profiles is thereby weakened; and the empirical specification adopted earlier is cast in doubt. We must therefore look at the robustness of the prediction.

The second question we must examine is that of the general shape described by the optimal work profile. Investment in human capital is thought to determine the lifetime profile of wage rates. The two are then presumed to combine multiplicatively to fashion the profile of earnings. Disentangling them again statistically, so that we may trace the influence paths and assess the importance of human capital and other factors, is a useful research task. To begin, we must try to glean from the theoretical arguments some testable hypotheses concerning how the wage and work profiles relate to one another--whether they are indeed concave functions, whether they have peaks within the relevant range, and if so, whether these peaks must occur in a given order.

The second section of this chapter draws in an informal way upon results of the utility-maximization approach. A simultaneous linear model of work and earnings is specified for estimation with the current data set. Results are reported and discussed in Chapter V.

THEORETICAL ANALYSIS

To date, there have been four major theoretical studies in which human-capital investment and labour supply appear simultaneously as endogenous variables. The earliest published, by Ghez and Becker,[5] uses traditional static utility maximization with discretely dated commodities to obtain the first-order conditions which characterize the solution to the individual's planning problem. This mode of analysis turns out to be sufficient to answer the two broad questions just posed; however, it does not provide a very rich understanding of the dynamic, processes involved. The other studies, by Blinder and Weiss,[6] by Heckman,[7] and by Ryder, Stafford, and Stephan,[8] employ control theoretic techniques to derive, within certain qualitative limits, the optimal profiles for investment, wages, and work. This survey will therefore emphasize the latter approach.

Since all four studies reach similar conclusions, it is not necessary--and it would in fact be redundant--to trace the mathematical details of each argument. Of greater interest are the particular assumptions which the various authors substitute for one another in deriving their results. The interchangeability of certain assumptions and the consistent necessity for others are the points to note in the following analysis. It is hoped that reducing the rather complex control theoretic studies to a single, uniform notation will also prove enlightening in itself.

Components of the Model

All the existing studies begin with an individual utility function such as

$$U = U(C, l) \quad , \qquad \ldots .(45)$$

defined over C, a composite Hicksian consumer good, and l, the quantity of leisure, measured as a proportion of the total time available.[9] Blinder and Weiss (B-W) assume strong separability, as do Ryder, Stafford, and Stephan (R-S-S), who specialize further by letting $U(C, l) = \ln(aC^{\theta_1}\ell^{\theta_2})$. Heckman ingeniously avoids separability by writing $U(C, l) = U(C, l \cdot H)$, where as before, H is the stock of human capital. The latter thus serves as an augmenting factor in the consumption of leisure. This specification is sufficient to produce determinate results, though it is not clear that it is a weaker postulate than separability. Heckman's illustrative findings and most of his comparative dynamic results stem from the CES case.

Apart from utility-producing leisure, the competing uses of time consist of work, denoted by m, and training, denoted by j. The time budget is simply

$$l + m + j = 1 \quad . \qquad \ldots .(46)$$

To connect this with the earlier analysis, let us define "market time" as $h \equiv m + j$. Then $k' = j/h$. R-S-S, along with Heckman, choose l and j as control variables for the optimization problem; B-W select h and k'. Since all are determined simultaneously, and since m is made dependent by (46), the choice is purely one of convenience. That of B-W meshes best with the previous discussion.

In addition to the time budget, the individual faces a lifetime expenditure constraint, which at any instant takes the form

$$\dot{A} = mwH + rA - C = (1-k')hE - rA - C \ , \qquad \ldots .(47)$$

where A represents nonhuman wealth, and $\dot{A}$, its time derivative. Recall that w and r signify the returns to human and nonhuman wealth respectively, and that E = wH is earning capacity. The price of consumption goods (the numeraire) has been set to unity.

B-W amend (47) in a subtle but important manner. In place of k' they write the negatively sloped, concave function g(k'). Whereas, Mincer utilizes $W/E = (1-k')$, they employ $W/E = g(k')$, with $g'(k') < 0$ and $g''(k') < 0$. B-W alertly point out that if the "earnings-investment frontier" g(k') were actually linear, as Mincer postulates, there would be no advantage to combining training and work. The individual could achieve any point on the frontier by dividing his time appropriately between pure training (k' = 1) and pure work (k' = 0). Since $g(k') > (1 - k')$ for $0 < k' < 1$, concavity makes on-the-job training uniquely profitable.[10]

It is worth noting in connection with (47) that there is no general restriction forcing A to assume nonnegative values. Individuals are free to borrow and to lend in a perfectly competitive financial market at the given rate r. Instead, one might think of implementing Becker's previously surveyed demand-and-supply model of human-capital investment by letting $r = r(A)$, with $r'(A) < 0$ for $A < 0$.[11] We shall observe shortly how this specification would complicate the analysis.

The final component of the present model is an equation describing the growth (and decay) of human assets. As in Chapter I, we may write:

$$Q_H = \alpha(k'hH)^{\mu} \quad ; \qquad \ldots .(19)$$

$$\dot{H} = Q_H - dH = \alpha(k'hH)^{\mu} - dH \quad , \qquad \ldots .(20)$$

except that, here, k' alone gives way to $k'h$ in recognition of the presumed variability in hours of potential investment time. R-S-S use precisely the foregoing specification. As we have seen, their assumption that $0 < \mu < 1$ ensures, with w constant, that the marginal cost of producing human capital is increasing. Heckman, on the other hand, manages with a general functional form, restricted only as to first and second partial derivatives and containing both time and purchased educational inputs.[12] B-W employ the special assumption that $\mu = 1$; accordingly,

$$\dot{H} = (\alpha k'h - d)H \quad . \qquad \ldots .(20')$$

They are able to proceed in this manner on account of $g(k')$. Concavity of the latter implies increasing marginal cost even though returns in production are constant. Since equilibrium and the time path of investment depend only on the shape of the marginal cost curve (given the shadow price of human capital), it does not appear that exchanging $\mu < 1$ for $g''(k') < 0$ has any effect on the generality of the results.

A formal statement of the control problem is now possible. It is to maximize[13]

$$\int_0^T e^{-\rho t} U(C, l)\,dt + B[A(T)] \quad ,$$

where ρ is the rate of time preference and $B[A(T)]$ is the (separable) utility of terminal assets, subject to (46), (47), and (20) and to[14]

$$h = m + j \geqq 0 \quad \text{and} \quad 0 \leqq k' \leqq 1 \quad , \qquad \ldots .(48)$$

given the initial conditions

$$H(0) = H_0 \geqq 0 \quad \text{and} \quad A(0) = A_0 \geqq 0 \quad . \qquad \ldots .(49)$$

The control variables are C, m, and j (or equivalently, C, h, and k'); the state variables are H and A.

Analysis

The Hamiltonian, based on the assumptions of B-W, may be written as follows:

$$J = e^{-\rho t}\Big\{U(C, l - h) + \lambda_A[g(k')hwH + rA - C] + \lambda_H[(\alpha k'h - d)H]\Big\} \quad . \qquad \ldots .(50)$$

As usual, λ_A and λ_H are shadow prices. The necessary conditions for an interior maximum take the following form:

$$\partial J/\partial C = 0 : \quad U_C = \lambda_A \qquad \ldots .(51)$$

$$\partial J/\partial h = 0 : \quad U_l = \lambda_A g(k')wH + \lambda_H \alpha k'H \qquad \ldots .(52)$$

$$\partial J/\partial k' = 0 : \quad 0 = \lambda_A g'(k')wH + \lambda_H \alpha H \qquad \ldots .(53)$$

$$\partial J/\partial A = -(d/dt)(\lambda_A e^{-\rho t}) : \dot{\lambda}_A/\lambda_A = \rho - r \qquad \ldots .(54)$$

$$\partial J/\partial H = -(d/dt)(\lambda_H e^{-\rho t}) : \dot{\lambda}_H/\lambda_H = \rho + d$$
$$- g(k')hw\lambda_A/\lambda_H - \alpha k'h \qquad \ldots .(55)$$

$$(\text{transversality}) : \lambda_H(T)H(T) = 0 \qquad \ldots .(56)$$

$$(\text{transversality}) : \lambda_A(T) = B'[A(T)]. \qquad \ldots .(57)$$

These conditions hold as a set wherever $h > 0$ and $0 < k' < 1$.

However, as we found in the case of the (Ben-Porath) income-maximization model, boundary solutions occur very readily, portraying familiar stages in the typical life cycle. Making leisure endogenous increases the possible number of such stages from two to four, namely: (I) "schooling" ($h > 0$, $k' = 1$); (II) "training" ($h > 0$, $0 < \underline{k}' < 0$); (III) "work" ($h > 0$, $k' = 0$); (IV) "retirement" ($h = 0$, k' arbitrary).

Since the data set utilized by the present study samples only from the population of individuals in stages II and III, this review will ignore the other phases of the optimal plan.[15]

Before we examine the profiles of work and investment implied by (51)-(57), it is worth pausing briefly to confirm the economic interpretation of these conditions. Equation (51) merely demands that the marginal utility of goods be set equal to their shadow price at each instant; (57) imposes the same requirement on the terminal stock. Equations (51) and (54) together imply the well-known life-cycle result that consumption falls, remains constant, or rises according to whether $\rho \gtrless r$.[16] Equation (52) states that the marginal cost of nonleisure activity (U_l) equals, first, the benefit in the form of real earnings ($\lambda_A g(k')wH$) and, second, the benefit in the form of increased human capital, or future earnings ($\lambda_H \alpha k'H$). If $k' = 0$ (stage III), the marginal rate of substitution between goods and leisure, U_l/U_C, simply equals the real wage, wH, just as in the static analysis; but otherwise, $U_l/U_C > wH$. Equation (53) requires that the individual allocate his market time in such a way that the marginal input cost in foregone earnings ($-\lambda_A g'(k')wH$) equals the marginal present and future benefit of increased earning potential ($\lambda_H \alpha H$).

It is also convenient at this point to note the effect of making r depend on A. Only (54) is altered: r is replaced by $r(A) + r'(A)A$. The change is nevertheless crucial, as it makes the evolution of the shadow price a function of the state variable. This situation greatly complicates the ensuing analysis, and it is not known whether all of the main conclusions stand. Based on Heckman's comparative dynamic results for changes in an exogenous rate of interest,[17] one might risk a guess that the principal effect would be to flatten the wage profile; however, nothing more is clearly apparent.

The other alternative assumptions--those concerning utility and the production of human capital--yield significant, though manageable changes in the preceding set of first-order conditions.[18] To accommodate the differences, the three control-theoretic papers adopt divergent analytical strategies, together with some further restrictions on behaviour. The reasons in each case are most easily understood if we follow for a moment the derivation of B-W.

These authors study, among other things, the optimal trajectories in (k',h)-space. If one differentiates (52) logarithmically with respect to time, it is possible to show, using (53), (54), (55), and (20'), that

$$\dot{h}[-U_{ll} \;/ U_l] = \rho - (r + d)/(1 + \eta) \quad , \qquad \ldots .(58)$$

where $\eta \equiv -k'g'(k')/g(k')$ is the elasticity of $g(k')$. A similar operation performed on (53) yields, eventually,

$$\dot{k}[g''(k')/g'(k')] = r + d - \alpha k'h(1 + \eta)/\eta \quad . \qquad \ldots .(59)$$

These expressions define two stationary loci $\dot{h} = 0$ and $\dot{k}' = 0$. A third, $\dot{H} = 0$, may be obtained directly from (20').

All three curves are shown in Figure 1, reproduced (with the appropriate notational amendments) from B-W.[19] It is easy to verify by straightforward manipulation of (58), (59), and (20') that: (a) h = 0 is a vertical line at $\hat{k}'$ $(0 < \hat{k} < 1)$; (b) $\dot{k}' = 0$ rises monotonically from $[0, -g'(0)(r+d)/\alpha]$ to $[1, (r+d)/\alpha]$;[20] (c) $\dot{H} = 0$ is the rectangular hyperbola $H = (d/\alpha)(1/k')$; (d) the intersection of $\dot{h} = 0$ and $k' = 0$, namely $(\hat{k}', \hat{h})$, lies above $\dot{H} = 0$ if (but not only if) $r > \rho$ of the unit square, or in other words, on the boundary of stage I, where $k' = 1$. It would appear from the indicated motions that, unlike P, some trajectories may cycle about the point (k',h); but as B-W explain, such paths cannot arise. The reason provides considerable insight into the problem of formulating successfully a model of the present kind. Inspection of (58) and (59) reveals that (given the constants) k' and h depend only upon k' and h. To each point in (k',h)-space there corresponds a unique motion, defined by $[\dot{k}'(k',h), \dot{h}(k',h)]$. However, to attain the vertical axis $(k' = 0)$, as all trajectories eventually must, a cyclical path would have to cross itself at an angle, implying two different motions at the point of intersection.[21] This situation *could* arise without contradiction if either or both k' and h depended on the state or costate variables. Ensuring that they do not (and that we may consequently work with a two-dimensional phase diagram) is a matter for careful theorization.

It is clear from Figure 1 that the B-W model provides the hoped-for theoretical conclusions. First, the gross propensity to invest (k') declines monotonically throughout stage I and is therefore nonincreasing over the whole life cycle. Second, the supply of market

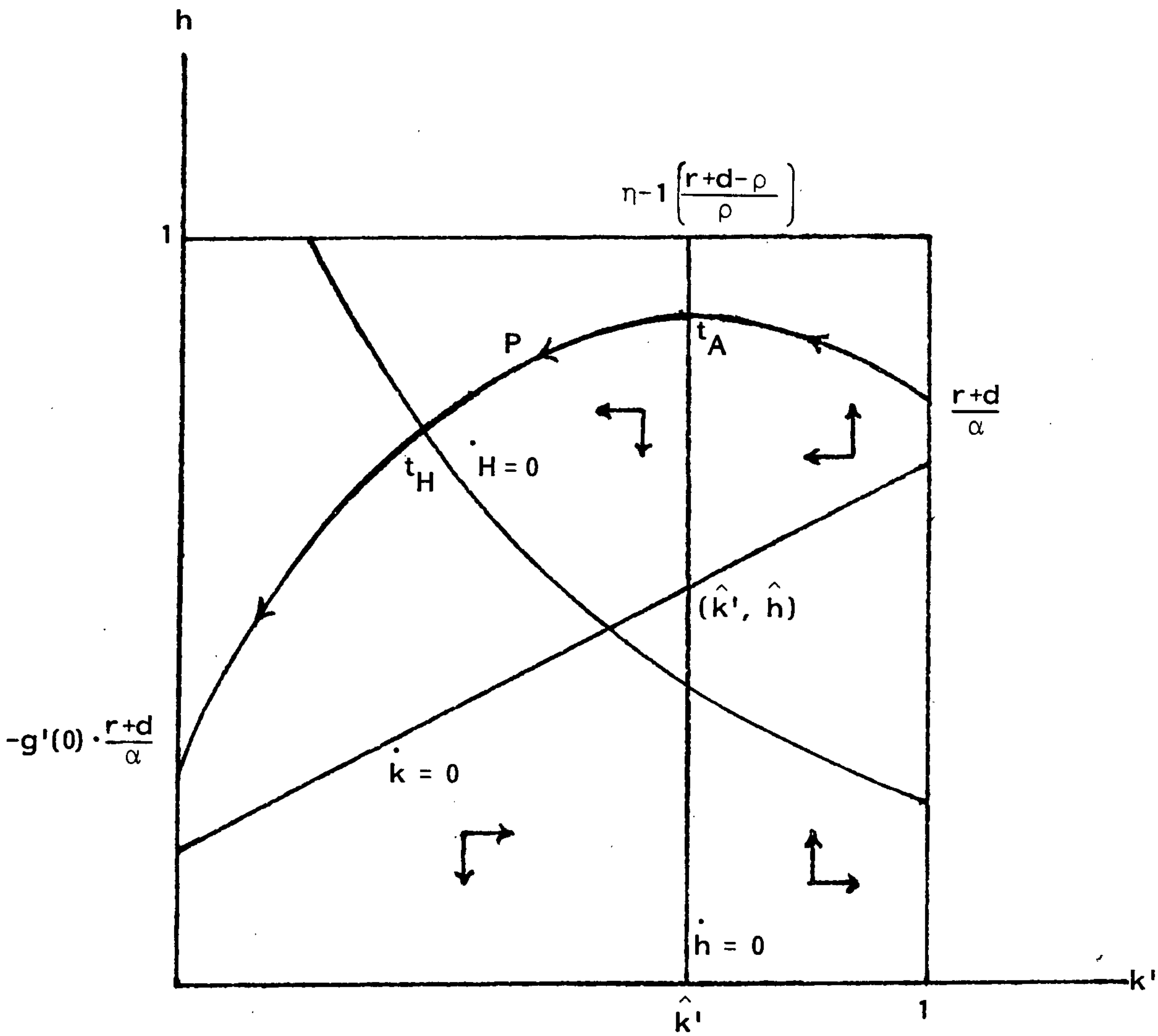

Figure 1 Phase diagram in (k', h)-space.

hours (h) rises to a peak at t_h and declines thereafter.[22] Third, if $r > \rho$ (a sufficient condition only), the peak in hours precedes t_H, the peak in human capital, which as we know, precedes the peak in measured earnings, whenever $d > 0$. These are the restrictions which, at a minimum, any empirical model must test.

As noted, the other two studies derive similar results by alternative means. Being unable to eliminate the unwanted state variables H, Heckman eschews the phase-diagrammatic approach in favour of solving the first-order conditions to obtain the demand functions for goods, effective leisure (lH), and investment (jH). Despite specializing the utility and production functions to the CES form, he cannot rule out locally increasing investment time except by means of the auxiliary assumption that depreciation is "small." Comparative dynamic investigation of changes initial wealth (human and nonhuman), the rate of interest, depreciation, ability, and taxes furnishes some interesting hypotheses, but apparently none which the author is able to test with the data at hand.

R-S-S are also faced with the presence of the state variable H on account of their nonlinear production specification. They proceed by letting $\rho = r = 0$. It is evident from Equation (54) that in this special case λ_A, is constant. Therefore, it is possible to draw a two dimensional phase diagram in (H, λ_H)-space and to deduce from it the behaviour of all the control variables. It turns out that h reaches its peak at the same time as H, though again, before the peak in measured earnings. As in Heckman, j cannot be shown to decline monotonically.

This result is not, of course, inconsistent with the B-W conclusion, stated in terms of k'. Since $j = k'h$, we have $\dot{j} = \dot{k}'h + k'\dot{h}$. The first term is always negative; the second is positive or negative according to whether $\dot{h} \gtrless 0$. Thus, even though the proportion of market time devoted to investment is unambiguously falling, investment time itself may be rising if total market time is increasing rapidly enough.

In summary, the theoretical analysis tends to weaken the human-capital interpretation of concave wage and earnings profiles by admitting the possibility of rising investment at some points in the life cycle. The analysis supports an empirical model which makes hours a peaked, concave function of age. Though certain comparative dynamic results have been adduced under strong assumptions, these predictions do not yield very readily to testing with cross-section data.

AN EMPIRICAL MODEL

This section introduces a simultaneous linear model of wages and hours which is simple enough to be estimated with the current data set. Though the model is incapable of settling all outstanding issues and is not conventionally rigourous in the sense of being derived from standard, known utility and production functions, it does appear to capture the most important measurable factors affecting individual decisions.[24]

Structural Equations

The model consists of an identity and two behavioural relationships:

$$Y = W_a h \quad \text{or} \quad \ln Y = \ln W_a + \ln h \qquad \ldots .(60)$$

$$W_a = e^{X'\beta} h^{\theta} u_1 \quad \text{or} \quad \ln W_a = X'\beta + \theta \ln h + \ln u_1 \qquad \ldots .(61)$$

$$h = e^{Z'\gamma} W_m^{\delta} u_2 \quad \text{or} \quad \ln h = Z'\gamma + \delta \ln W_m + \ln u_2 \qquad \ldots .(62)$$

For each individual (subscript suppressed), annual employment earnings, Y, are the product of the average hourly wage before tax, W_a, and the number of hours worked, h. The average wage depends, first of all, upon h. Conversely, h depends upon another row vector of determinants, Z', which may have elements in common with X', and upon the marginal after-tax wage, W_m. Among the remaining symbols, u_1 and u_2 are stochastic terms; β, γ, θ, and δ are vector and scalar constants, as the context indicates.

Observe that if we substitute (61) into (60), the result is

$$y = e^{X'\beta} h^{(1+\theta)} u_1 \quad . \qquad \ldots .(63)$$

Then, if τ represents the marginal tax rate on earnings (assumed for the moment to be constant), the marginal after-tax wage must be given by

$$W_m = (1 - \tau) \cdot \partial Y / \partial h$$

$$= (1 - \tau)(1 + \theta) e^{X'\beta} h^{\theta} u_1$$

$$= (1 - \tau)(1 + \theta) W_a$$

$$= (1 - \tau)(1 + \theta) \cdot Y/h \quad .$$

Substituting into the logarithmic version of (62) yields

$$\ln h = Z'\gamma + \delta \ln [(1 - \tau)(1 + \theta) Y/h] + \ln u_2$$

$$= Z'\gamma + \delta \ln (1 - \tau) + \delta \ln(1 + \theta) + \delta \ln Y - \delta \ln h + \ln u_2 \quad .$$

Solving the latter for ln h and taking the logarithm of (63), one finally obtains a pair of estimable equations:

$$\ln Y = X'\beta + (1 + \theta) \ln h + \ln u_1 \qquad \ldots .(64)$$

$$\ln h = \frac{1}{1 + \delta} Z'\gamma + \frac{\delta}{1 + \delta} \ln (1 - \tau) + \frac{\delta}{1 + \delta} \ln Y + \frac{\ln u_2}{1 + \delta} \quad . \qquad \ldots .(65)$$

These form the basis for the work reported in Chapter V.

Further Comment and Definition

Now that the general outlines of the model are clear, it is possible to discuss the specification in some detail. The preceding equations contain a number of distinct hypotheses which require amplification, and it is of course essential to define the constituents of X' and Z'.

The first thing to note is that although (61) and (62) are "structural" equations from the standpoint of the model, they are not the structural equations one might conventionally use to segregate supply and demand in the labour market. Here, supply and demand factors presumably mingle in forming the respective lists (X' and Z') of exogenous variables. Therefore, it is not immediately clear whether one should take as an endogenous variable the price firms pay for labour (W_a) or the price individuals ultimately receive for it (W_m). Equation (61) employs W_a, making X' and h the determinants of *average* gross worker productivity.[25] Since schooling and experience (elements of X') are still taken to be exogenous, or at the very least predetermined, the fact that individuals in a given cross section might once have considered W_m in formulating their investment plans is not necessarily relevant. Equation (62) incorporates the standard labour-supply assumption that individuals respond to the *marginal* net wage.

Although the insertion of W_m in (62) may appear unremarkable, its use does require some justification in a life-cycle context. When work and investment are planned simultaneously, the individual does not (except in stage III) equate his marginal rate of substitution between goods and leisure to the net wage, as the static theory implies. Moreover, since the lifetime profile of W_m is known ex ante, the effect of this variable upon time worked *at any given moment* is not of the standard causal variety. The two must harmonize in the optimal plan; that is all. Accordingly, one might think of replacing W_m with some function of age or experience which depicts the outcome of the initial planning decision.

The explicit inclusion of W_m is nevertheless indicated on a number of grounds. In the first place, W_m may characterize the optimal plan more accurately than a purely exogenous function of the sort just mentioned. There is no harm in using the endogenous variable so long as we are not mislead into making unwarranted inferences concerning static income and substitution effects. Secondly, though work and investment may evolve together in a planned way during the period of on-the-job training, labour supply may respond causally to that component of the net wage which is the result of predetermined schooling and the initial endowment of human capital. Finally, one must concede that in the real world the wage rate will be subject to unforeseen disturbances. The individual will presumably want to adjust his work effort to these, much as the static theory suggests.[26]

The use of h as a determinant of W_a likewise appears justified on several counts. Moonlighting and overtime are the two which come most quickly to mind.[27] Both affect the average wage by altering the remuneration earned on succeeding increments of work. If secondary employment pays less per hour than primary, moonlighting will influence

θ toward the negative. The existence of an overtime premium will deflect it toward the positive. If h acts as a proxy for various motivational, ability, and environmental factors which serve as common determinants of wages and employment, there is further reason to expect that θ will be nonzero. Since most of the personal factors one can name would appear to operate upon wages and employment in the same direction, it seems likely on this ground that $\theta > 0$. However, if the labour market actually works in an oppressive manner, heaping long hours upon the poorly paid (and conversely, favouring the best paid with abundant leisure), then it may turn out, as in Chapter III, that $\theta < 0$. The same may occur, as suggested earlier, if seasonal workers obtain high wages to compensate for limited hours. One cannot predict, but it is certainly important to estimate, the sign and the significance of this parameter.[28]

Estimation, by means of (64) and (65), is relatively straightforward once the elements of X' and Z' have been defined. Since the approach taken here is to a certain degree experimental, it would be inappropriate to specify the exact composition of these vectors in advance. However, it is useful at this point to discuss the most prominent candidates for inclusion.

With regard to X', only a brief comment is required. Obviously, one would wish to define this vector in terms of the variables found significant in the single-equation estimates of Chapter III. Though all are potentially admissible as elements of X', emphasis will be given in Chapter V to the human-capital variables appearing in the orthodox earnings function. With X' restricted in this way, assessment of the latter in light of the simultaneous-equation estimates is greatly facilitated. Variables will nevertheless be added to X', as they were to the orthodox earnings function--in the present case, to distinguish their separate influences upon wage rates and hours of work.

With regard to Z', more needs to be said than in the preceding instance, since we have not elsewhere considered the likely determinants of hours worked.[29] It should be clear, even so, that two essential components of Z' must be age and schooling. These variables are key factors in the present inquiry, and their use in an equation like (65) is well established in the literature. Age will surely affect hours worked if the preceding life-cycle theory is valid. To test its prediction of peakedness in the age-hours profile, we shall let Z' include both age and age squared.[30] Schooling may affect realized hours in a number of ways: by determining the sort of job (high-unemployment or low-unemployment) that a worker may hold, by determining the efficiency of job search, by indicating worker quality to prospective employers,by conditioning the susceptibility to layoff.[31] It is of considerable interest to compare the effect schooling may have upon earnings by way of hours with the effect it evidently has upon earnings by way of wage rates. Including the variable in both X' and Z' should furnish the desired information.

Other plausible components of Z' are family status, ethnic group, industry and occupation, and place of residence. The first variable, consisting in detail of headship and marital status, is almost universal in the literature, though it commonly appears not as a regressor, but as a criterion with which to define subsamples for separate estimation. Ethnic group may affect hours through discrimination and through various culturally determined traits, as we have already inferred from the single-equation results. Industry and occupation are reasonable proxies for the employment characteristics of the jobs thus described. Residence is another proxy for employment conditions, which vary considerably across regions and no doubt influence the hours of work realized by the typical individual.

A final and very important component of Z' arises on stricter theoretical grounds. It is routine in the static analysis of labour supply to include in the resulting empirical equations an independent variable to portray the nonemployment income of the individual or family. The estimated coefficient of this variable then measures the static income effect. Such income effects also occur in the life-cycle model, though they are presumably spread over the whole planning period. In any event, they may be accounted for in the standard way. At the same time, it is necessary to relax the assumption that the marginal tax rate τ is constant. These two theoretical considerations combine to suggest a new income variable.

Its definition is illustrated with the help of Figure 2. This shows, in leisure-income space, the before-tax budget constraint BB' and the after-tax budget constraint AA' of an individual whose gross wage is

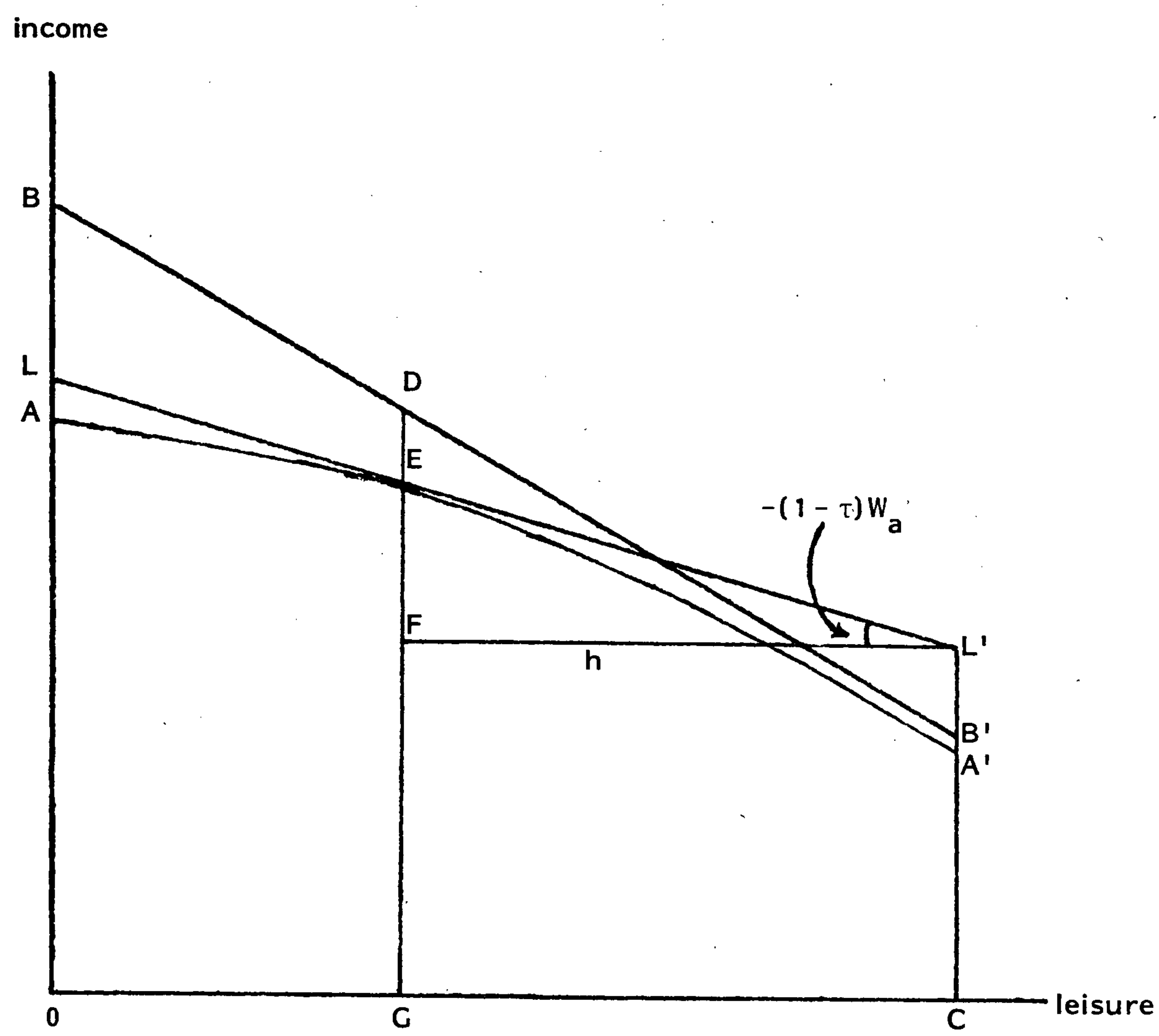

Figure 2 Linearization of the budget constraint

constant. The curvature of AA' (smoothed for purposes of illustration) reflects the progressivity of the tax system. Following Hall's procedure,[32] one may linearize the after-tax budget constraint at the observed equilibrium point E. The individual may then be assumed to behave as if he were facing LL', which (given the wage rate and the level of non-employment income B'C) is uniquely determined by the slope $(1-\tau)W_a = (1-\tau)Y/h$ and the zero-work intercept L'C.[33] The latter is given geometrically by DG - DE - EF, where $EF = h \cdot (1-\tau)Y/h = (1-\tau)Y$ and where DG represents total income and DE, total taxes. Knowing all these quantities, one may compute L'C for each individual and obtain the desired variable to include in Z'. Earlier, in Table 3, this variable was labelled INCOTH.

It must be noted that the foregoing procedure is at best appropriate only when the individual's gross wage is constant, as shown (or when equilibrium occurs only on the right-most segment of a piece-wise linear budget constraint). Otherwise, the slope of the budget constraint will be $(1-\tau)\cdot(1+\theta)W_a$, where θ is not known in advance. If nonzero values of θ arise purely through the correlation of wages and hours over the cross section (that is, among different jobs), then of course, the procedure remains ostensibly valid. However, if nonzero values arise for each individual (that is, within the terms of the job or jobs held), there will be errors in the calculation of INCOTH. It thus appears that the Hall procedure is capable of digesting only a certain degree of non-linearity in the budget constraint. Other difficulties associated with the approach--ones of an econometric nature--will be reviewed in Chapter V.

Meanwhile, a final point to consider in defining the intercept term is whether one should use merely the individual's own property earnings or the sum of these and the total income of all other family members.[34] Notwithstanding recent analyses of family labour supply,[35] it was found that in the present, rather heterogeneous sample "own property income" performed slightly better than "other family income" as a predictor of hours when (65) was subjected to preliminary examination by ordinary least squares.[36] Since the present purpose in estimating (65) is not to investigate labour supply as such, but rather to obtain the best instruments for use in system estimates focussing on (64), it was decided to adopt the narrower income concept--which accounts for the definition of INCOTH.

Although an equation like (64) is commonly referred to as a labour-*supply* function, this interpretation depends on a number of strong, usually implicit assumptions concerning the nature of demand and the relative variability of demand and supply. Whether or not one might actually identify a supply function in estimating (64) is difficult to say with complete confidence.[37] The present study takes an agnostic, empiricist approach to this question. Partly as a result, there were few constraints but also little guidance in selecting a functional form. The double-logarithmic or constant-elasticity form ultimately chosen to relate hours and the wage rate is highly convenient, though somewhat novel from the standpoint of the labour-supply literature, which has leaned toward the double-absolute (variable-elasticity) specification.[38] Regardless of whether the double-logarithmic form provides a convincing a priori description of labour supply, it appears to perform reasonably well as a predictor of hours. Some ordinary-least-squares estimates documenting this performance, along with that of the listed independent variables, are presented for inspection and comparison in the appendix which follows.

APPENDIX IV

ORDINARY-LEAST-SQUARES ESTIMATES OF WORKING HOURS[1]

(H1)	WTIME	=	-1.4522	+	.0675 AGE	-	.0008 ASQ		
			(47.0)		(43.5)		(43.4)		
	R^2	=	.077		number of observations	=	22,682		
(H2)	WTIME	=	-1.3585	+	.0236 AGE	-	.0003 ASQ	+	.4018 ZINC
			(44.0)		(16.2)		(15.6)		(70.3)
			-.1069 XINCOTHDI	+	.1640 DI				
			(22.4)		(7.98)				
	R^2	=	.309		number of observations	=	22,682		
(H3)	WTIME	=	-1.2801	+	.0225 AGE	-	.0003 ASQ	+	.4068 ZINC
			(37.7)		(15.3)		(15.0)		(70.4)
			-.1024 XINCOTHDI	+	.1629 DI		-.0055 S		
			(21.2)		(7.93)		(5.52)		
	R^2	=	.310		number of observations	=	22,682		
(H4)	WTIME	=	-1.2454	+	.0202 AGE	-	.0002 ASQ	+	.4373 ZINC
			(34.4)		(14.1)		(14.9)		(75.5)
			-.0919 XINCOTHDI	+	.1734 DI	-	.0033 S	+	.0292 GEO1
			(18.9)		(8.71)		(2.89)		(2.35)
			-.0318 GEO2	+	.0341 GEO4	-	.0081 GEO5	-	.0894 GEO6
			(4.03)		(2.87)		(0.65)		(8.31)
			.0256 TYPE	+	.3731 IND1	-	.0946 IND2	-	.0746 IND3
			(3.73)		(15.1)		(3.34)		(1.65)
			-.0024 IND4	-	.1072 IND6	-	.0152 IND7	+	.0739 IND8
			(0.11)		(7.72)		(1.31)		(7.09)
			-.0058 IND9	-	.0076 IND10	+	.0817 MAJ	-	.1079 OC1
			(0.33)		(0.64)		(6.92)		(5.64)
			-.1033 OC2	-	.1657 OC3	-	.0952 OC4	-	.0436 OC5
			(5.14)		(7.20)		(3.61)		(2.48)
			-.0015 OC6	-	.1092 OC8	-	.0351 OC9	-	.1179 OC10
			(0.84)		(4.62)		(2.26)		(6.51)
			-.0312 OC11	-	.0790 OC12				
			(1.71)		(4.90)				
	R^2	=	.359		number of observations	=	22,682		
(H5)	WTIME	=	-1.1766	+	.0161 AGE	-	.0002 ASQ	+	.4980 ZINC
			(31.9)		(10.6)		(11.8)		(57.9)
			-.0887 XINCOTHDI	+	.0846 DI	-	.0028 S		
			(17.4)		(4.01)		(2.43)		
			+ • • • (GEO, TYPE, IND, MAJ, OC) • • •						
			+.1523 HEAD	-	.0010 FAMSIZ	+	.0460 USMAR		
			(6.00)		(0.56)		(2.12)		
			-.0106 ETH2	-	.0182 ETH3	-	.0063 ETH4		
			(1.37)		(1.37)		(0.18)		
			-.0059 ETH5	-	.3511 ETH6				
			(0.23)		(9.07)				
	R^2	=	.367		number of observations	=	22,682		

[1]Figures in parentheses are t ratios, written in absolute terms.

NOTES

CHAPTER IV

[1]In the perfectly competitive labour market implicitly assumed, the two are of course identical.

[2]See, for example, Gary S. Becker, "The Allocation of Time over the Life Cycle," in Gilbert R. Ghez and Gary S. Becker, The Allocation of Time and Goods over the Life Cycle (New York: Columbia University Press for the National Bureau of Economic Research, 1975).

[3]If time were not an inelastically supplied resource, independence might still be maintained, since the quantity used in consumption would then not affect the price or the quantity available for use in investment. Fixity of the time endowment, rather than multiple use, is therefore the key element of the problem.

[4]It is possible, of course, to restrict the underlying utility function in such a way that the simpler model will yet suffice. Suppose that the individual is initially in equilibrium, equating the marginal rate of substitution between goods and leisure to the net wage. If he then decides to allocate some nonleisure time to investment, the net wage will fall in the current period and rise thereafter. If equilibrium is to be restored without upsetting the investment calculation, labour supply must not change. The utility function must render the demand for leisure perfectly inelastic. Needless to say, this is a very strong requirement.

[5]Op. cit.

[6]Alan S. Blinder and Yoram Weiss, "Human Capital and Labor Supply: A Synthesis," Journal of Political Economy, LXXXIV (June, 1976), 449-472.

[7]James J. Heckman, "A Life-Cycle Model of Earnings, Learning, and Consumption," Journal of Political Economy, LXXXIV (August, 1976), S11-S44.

[8] Harl E. Ryder, Frank P. Stafford, and Paula E. Stephan, "Labor, Leisure, and Training over the Life Cycle," International Economic Review, XVII (October, 1976), 651-674.

[9] Though functional notation has been suppressed, all variables implicitly depend on time.

[10] Note that, by hypothesis, $g(1) = 0$ and $g(0) = 1$.

[11] If the holders of large positive asset portfolios obtain the highest net returns, $r(A)$ might in fact by U-shaped, with $r'(A) > 0$ for $A > 0$. A discontinuity at $A = 0$ is certainly to be expected.

[12] His specification is $Q_H = F(bk'hH, D)$, where b is a constant quickly set to equal unity. The presence of b avoids the particular neutrality assumption implicit in making H the augmenting factor in both the utility and the production function.

[13] Note that, here, T designates the termination of the optimal plan, not the point of zero net investment, as in the discussion of Mincer.

[14] One may either add $I > 0$ and $A(T) \geqq 0$ or restrict the utility function so that the respective marginal utilities become arbitrarily great at zero. This ensures nonnegativity in any optimum.

[15] For a complete statement of the first-order conditions see B-W, op. cit., p. 457.

[16] Cf. Ghez and Becker, op. cit., who find that the profile of consumption imitates the profile of wage rates. This conclusion stems from the authors' adherence to Becker's theory of time allocation, which suggests that individuals substitute market goods for leisure in household production as the wage rate rises.

[17] Op. cit., pp. 526-527.

[18] Heckman's utility function adds the factor H to the left-hand side of (52), making it possible to concel H completely, but contributes the term $-(1 - h)U_\ell/\lambda_H$ to (55). R-S-S replace U_ℓ with the special form θ_2/ℓ. The final term in (52) becomes $\lambda_H \mu Q_H / h$, amd (53) becomes $0 = \lambda_A g'(k')hwH + \lambda_H \mu Q_H / k'$. In (55), $\mu Q_H / H$ replaces $\alpha k'h$. All authors except B-W assume $g(k') = (1 - k')$, whence $g'(k') = -1$.

[19] *Op. cit.*, p. 464, Figure 4.

[20] As the reader may verify, additional properties of $\dot{k}' = 0$ depend on the third and higher derivatives of $g(k')$, which are unspecified. B-W choose tacitly to depict the locus as a straight line.

[21] B-W do not mention the apparent possibility that $(\hat{k}', \hat{h})$ might be a stable focus. This is ruled out by the transversality condition (56). *Op. cit.*, p. 465.

[22] It continues to decline in stage III (pure work) if and only if $r + d - \rho > 0$, which B-W take to be the "leading case." *Op. cit.*, p. 463.

[23] Heckman, *op. cit.*, p. 518.

[24] Blinder uses a similar model for purposes of argument but does not pursue its implementation. See "On Dogmatism in Human Capital Theory," pp. 16-17.

[25] This is not to suggest that individual firms ignore marginal calculations, only that there is an empirical market relationship between W_a and the variables named.

[26] An essentially analogous argument relating consumption and earnings appears in Keizo Nagatani, "Life-Cycle-Saving: Theory and Fact," *American Economic Review*, LXII (June, 1972), 344-353.

[27] On these topics see: Robert Shishko and Bernard Rostker, "The Economics of Multiple Job Holding," *American Economic Review*, LXVI (June, 1976), 298-308; Yoram Barzel, "The Determination of Daily Hours and Wages," *Quarterly Journal of Economics*, LXXXVII (May, 1973), 220-238.

[28] For additional discussion and empirical analysis based on a sample of female workers, see Harvey S. Rosen, "Taxes in a Labor Supply Model with Joint Wage-Hours Determination," *Econometrica*, XLIV (May, 1976), 485-508.

[29] Empirical studies of labour supply investigate a number of factors, generally viewed as representing tastes or external constraints. See, for example, Marvin Kosters, "Effects of an Income Tax on Labor

Supply," in The Taxation of Income from Capital, edited by Arnold C. Harberger and Martin J. Bailey (Washington: The Brookings Institution, 1969); Sherwin Rosen and Finis Welch, "Labor Supply and Income Redistribution," Review of Economics and Statistics, LIII (August, 1971), 278-282; the collection of articles appearing in Income Maintenance and Labor Supply, edited by Glen G. Cain and Harold W. Watts (Chicago: Rand McNally College Publishing Company, 1973); Julie Da Vanzo, Dennis DeTray, and David H. Greenberg, "The Sensitivity of Male Labor Supply Estimates to Choice of Assumptions," Review of Economics and Statistics, LVIII (August, 1976), 313-325.

[30] Cf. Orley Ashenfelter and James Heckman, "Estimating Labor-Supply Functions" in Cain and Watts, op. cit.

[31] For more discussion see Farrell E. Bloch and Sharon P. Smith, "Human Capital and Labor Market Employment," Journal of Human Resources, XII (Fall, 1977), 550-560.

[32] Robert E. Hall, "Wages, Income, and Hours of Work in the U.S. Labor Force," in Cain and Watts, op. cit., pp. 118-121. For some additional discussion see W. Erwin Diewert, "Choice on Labor Markets and the Theory of Allocation of Time." (Unpublished discussion paper, Canada, Department of Manpower and Immigration, 1971).

[33] Hall actually uses the zero-leisure intercept LO.

[34] The present data do not allow a further subdivision of other family members' income into employment and nonemployment components. At best, one might apply the individual-utility-family-constraint model of Jane H. Leuthold, "An Empirical Study of Formula become Transfered and the Work Decision of the Poor," Journal of Human Resources, III (Summer, 1968), 312-323.

[35] See, for example: Reuben Gronau, "The Intrafamily Allocation of Time: The Value of the Housewives' Time," American Economic Review, LXIII (September, 1973), 634-651; Orley Ashenfelter and James Heckman, "The Estimation of Income and Substitution Effects in a Model of Family Labor Supply," Econometrica, XLII (January, 1974), 73-85.

[36] Greater measurement error in the latter (originally provided in class intervals), the inclusion of heads and nonheads of families, and the failure to distinguish between the property and nonproperty income of family members may have contributed to this result.

[37]This neglected problem has been discussed by A.C. Raynor, "On the Identification of the Supply Curve of Working Hours," Oxford Economic Papers, XXI (July, 1969), 293-298.

[38]An exception is the time-series expenditure study of Michael Abbot and Orley Ashenfelter, "Labor Supply, Commodity Demand and the Allocation of Time," Review of Economic Studies, XLIII (October, 1976), 389-412. Cross-section precedents include: Lee Lillard, "Estimation of Permanent and Transitory Responses in Panel Data: A Dynamic Labor Supply Model." (Unpublished report, Santa Monica: RAND Corporation, 1977); Gary Burtless and Jerry A. Hausman, "The Effect of Taxation on Labor Supply: Evaluating the Gary Negative Income Tax Experiment," Journal of Political Economy, LXXXVI (December, 1978), 1103-1130.

CHAPTER V

EARNINGS AND HOURS: SIMULTANEOUS-EQUATION ESTIMATES FOR CANADA

The preceding chapter develops a simplified, linear version of the earnings-and-hours model. Though we have dealt at some length with the economic content of the proposed specification, nothing has yet been said regarding the econometric assumptions and procedures needed to implement it. Accordingly, the first section of this chapter discusses estimation. The second reports results and offers an analysis.

ESTIMATION PROCEDURE

Before we may consider the choice of a particular econometric technique for estimating the two-equation model, it is necessary to define the stochastic framework. So far, no restrictions have been placed upon the disturbances appearing in (64) and (65). For convenience, these equations are restated here as a system in "stacked" matrix form:[1]

$$\begin{pmatrix} \ln Y \\ \ln h \end{pmatrix} = \begin{pmatrix} X & 0 & 0 \\ 0 & Z & \ln(1-\tau) \end{pmatrix} \cdot \begin{pmatrix} \beta \\ \gamma \cdot 1/(1+\delta) \\ \delta/(1+\delta) \end{pmatrix}$$

$$+ \begin{bmatrix} \ln h \cdot (1+\theta) \\ \ln Y \cdot \delta/(1+\delta) \end{bmatrix} \begin{bmatrix} \tilde{u}_1 \\ \tilde{u}_2 \end{bmatrix} . \qquad \ldots .(66)$$

Since δ is a constant, there is no harm in treating $\tilde{u}_2 \equiv (\ln u_2)/(1+\delta)$ as an ordinary random error, like $\tilde{u}_1 \equiv \ln u_1$.

It is reasonable to assume the following:

$$E(\tilde{u}_{1i}) = E(\tilde{u}_{2i}) = 0$$

$$E(\tilde{u}_{1i}\tilde{u}_{1i}) = \sigma_{YY} \qquad E(\tilde{u}_{2i}\tilde{u}_{2i}) = \sigma_{hh} \qquad i, j = 1, 2, \cdots, N$$

$$E(\tilde{u}_{1i}\tilde{u}_{2i}) = \sigma_{Yh}$$

$$E(\tilde{u}_{1i}\tilde{u}_{1j}) = E(\tilde{u}_{2i}\tilde{u}_{2j}) = E(\tilde{u}_{1i}\tilde{u}_{2j}) = 0 \qquad i \neq j . \qquad \ldots .(67)$$

Within each structural equation individual errors are homoskedastic; in general, however, the common variances are not the same across equations ($\sigma_{YY} \neq \sigma_{hh}$). For each individual the covariances across equations are also uniform (equalling σ_{Yh}), but their common value need not be zero. Since omitted variables--factors special to the individual or to his particular environment--may affect both earnings (via the wage rate) and hours, one cannot assume that $\tilde{u}_{1i}$ and $\tilde{u}_{2i}$ will be uncorrelated. One can safely assume that between all given *pairs* of individuals the covariances within and across equation will be zero. If we let

$U' = [u_1' \, u_2']$, the variance-covariance matrix of structural disturbances consistent with (67) may be written as follows:

$$E(UU') = \sum \otimes I_N \ , \quad \text{where } \sum = \begin{pmatrix} \sigma_{YY} & \sigma_{Yh} \\ \sigma_{Yh} & \sigma_{hh} \end{pmatrix} . \qquad \ldots . (68)$$

That is, $E(UU')$ consists of four $N \times N$ submatrices, each with the corresponding element of $\sum$ down the main diagonal and zeros elsewhere. For purposes of hypothesis testing we shall want to assume that $\tilde{u}_1$ and $\tilde{u}_2$ are normally distributed.

If one could ignore σ_{Yh}, it would be possible to obtain consistent, asymptotically efficient estimates of (66) using an instrumental-variable or two-stage least-squares regression procedure, equation by equation. However, the strong probability of a significant cross-equation covariance means that such methods are unlikely to be asymptotically efficient in the present case. Three-stage least squares (3SLS) would therefore seem to be a logical choice.[2] This estimator is both consistent and asymptotically efficient under given stochastic assumptions. Though it may differ numerically in finite samples from the full-information maximum-likelihood estimator, the two have the same asymptotic distribution.[3]

In carrying out the 3SLS procedure, one uses, in effect (though not computationally), the residuals from the second-stage (instrumental-variable) regression to form a consistent estimate of $\sum$. "Stage three" then amounts to performing general least squares (GLS) on the stage-two variables. Since the result, in general, is a new set of consistently estimated residuals, it is possible to repeat the GLS procedure until the regression coefficients cease changing. This technique, known as iterative 3SLS, cannot be shown to increase asymptotic efficiency[4] but may appear to some less arbitrary than stopping after one round. The iterative version of 3SLS is not adopted here, essentially on pragmatic grounds: estimates obtained by this means appear unrealistic in comparison with those obtained by ordinary 3SLS. As evidence, some iterative estimates are displayed in Appendix V.

Though it might seem that we are now in a position to examine results, the fact is that several important econometric issues remain to be discussed. These have to do with (1) the endogeneity of the tax rate, (2) the nature of the time-worked variable, and (3) identification. Let us consider each problem in turn.

First of all, because the marginal tax rate (τ) depends directly upon earnings, and therefore indirectly upon time worked, it is clearly an endogenous variable. The Hall procedure, described in Chapter IV, requires that we use the marginal tax rate in forming a slope and an intercept term, both of which are to appear on the right-hand side of any time-worked equation. In the notation of (66) the slope variable, obtained by combining terms, is $\ln(1-\tau)Y$; the intercept variable is a constituent of Z. Empirically, ZINC has been defined to represent the former; INCOTH, the latter. Furthermore, as explained in Chapter IV, INCOTH is replaced in practice by the dummy-interaction pair DI and XINCOTHDI. Since all three variables--ZINC, DI, and XINCOTHDI--are endogenous, their use in the time-worked equation of (66) will presumably result in biased estimates unless further steps are taken. In short, though the Hall procedure achieves the mapping of individual equilibria, it is not unblemished econometrically.[5]

One way round the problem--an approach used here and elsewhere[6]--is to form instrumental-variable estimates of the endogenous income terms. This technique should yield consistent final estimates of the structural coefficients, but it is difficult to apply in the present circumstances on account of the nonlinearity in the tax schedule,[7] the very problem which leads to endogeneity in the first place. Nevertheless, ZINC and XINCOTHDI were subjected to the instrumental-variable treatment, the instruments being those exogenous variables needed to simulate the tax rate and those found important in explaining INC.[8] Among the instruments were, in particular, the quadratic terms SSQ and PSQ. One would hope that these terms might go some way towards approximating the expected nonlinearity of the predicting equations. Since dummy-variable strings comprise the remaining instruments, functional forms were not in any event acutely constrained.

The use of ZINC serves to inforce the hypothesized equality restriction on the coefficients of $\ln(1 - \tau)$ and $\ln Y$. Where this was undesirable, it was necessary to form separate instrumental-variable estimates of the two terms, represented empirically by TMARG and INC. The same exogenous variables were employed in each case.

The endogenous dummy variable DI was left "unpurged," owing to the computational expense involved and to its dichotomous nature, which prevents efficient estimation by linear least squares. This omission does not seem very serious, since DI is *not* equal to one only for those individuals who fall in the zero-tax bracket *and* have no property income.[9] Because the zero-tax bracket is relatively wide, DI is furthermore unlikely to change very often in response to the disturbances in the earnings equation; in other words, DI and these disturbances will not be highly correlated.[10] The endogeneity problem is therefore likely to be minimal.

We come now to the second econometric issue, that of the time-worked variable. The PUS data available for measuring time worked are, on the whole, rather disappointing. It was decided that WTIME, as opposed to WEEKS, should stand for the theoretical variable h, even though the latter produced slightly better fits in the ordinary least-squares (OLS) regressions. Whereas WTIME may take on thirty-five different values, WEEKS is limited to only five.[11] The former thus resembles more closely than the latter the continuous variable we have in mind. Estimation using WEEKS would appear more suited to one of the probability models, such as the multinomial logit.

Both WEEKS and WTIME constitute "limited dependent variables," but the problem with regard to WEEKS is undoubtedly the more severe. By definition, WEEKS must fall in the half-closed interval[12] $(0, 52]$, with many observations lying on the upper bound. WTIME must exceed zero; but apart from the limit imposed in practice by grouping, there is no firm upper bound within the normal range of experience. Though observations are likely to be relatively dense in the vicinity of 2,000 hours, some individuals will report working a much larger accumulation. Hence, the distribution of hours, and of the disturbance in any WTIME equation one might estimate, need not be truncated on the upper side to any noticeable degree. The problem of the zero bound will be ignored here.[13] Conclusions regarding hours worked will thus be of the "conditional" variety.

The final problem we have to consider is that of identification. There is no gain in applying 3SLS to a given equation of the system unless the *other* is overidentified.[14] That the earnings equation, expressed in the human-capital form, is overidentified should be obvious, since many variables to be used in explaining hours are excluded from it. That the hours equation will also be overidentified may not be so clear. The matter rests on the empirical use of age and experience.

On the basis of the life-cycle analysis presented in Chapter IV, and in the absence of arguments to the contrary, AGE and ASQ were used in the hours equation. The experience variables P and PSQ, which do not appear in the latter, continue on the right-hand side of the earnings equation. Their exclusion from the hours equation would appear to settle the issue of overidentification, but one must remember that in practice $P = AGE - S - 5.67$. Accordingly, $PSQ = P^2 = ASQ + SSQ - 2 \cdot AGE \cdot S - 11.34\ \text{age} - 11.34\ S + 32.15$. Therefore, to the extent that the hours equation is overidentified, it will be through the exclusion of the variables SSQ and AGE • S. Since these terms enter the human-capital earnings equation (through PSQ) with an equality restriction on their coefficients, identification will not be so strong, however, as in the usual, unrestricted case.

RESULTS

Tables 26 and 27 report estimates of the structural equations pertaining, respectively, to earnings and to hours. Equations with the same numeric digit in their reference codes were estimated simultaneously. Since the earnings equation was of primary interest, the specification of the hours equation was held constant--the one exception being in (MH2), where the equality restriction on the coefficients of $(1-\tau)$ and Y (TMARG and INC) was briefly relaxed. Experiments with the earnings equation involved the addition of SSQ, XSP, GEO, TYPE, IND, MAJ, IM, and ETH to the basic human-capital formulation.

Initial Findings

The basic formulation appears in (ME1) and (ME2). The most striking feature of these equations--or for that matter, of the entire set--is the dramatic rise in the coefficient of WTIME. The values displayed here are more than double the one obtained by OLS.[15] Qualitatively, this outcome tends to reverse the finding in Chapter III that earnings respond inelastically to a change in hours. Quantitatively,

TABLE 26

SIMULTANEOUS ESTIMATES:[a] EARNINGS

Right-Hand Variable	Equations[b] (dependent variable = INC)			
	(ME1)	(ME2)	(ME3)	(ME4)
Constant	1.0021 (34.1)	.9427 (34.0)	1.1652 (18.2)	1.3676 (35.6)
S	.0629 (36.6)	.0640 (37.1)	.0009 (1.26)	.0525 (24.1)
SSQ	- -	- -	.0025 (9.22)	- -
P	.0221 (15.5)	.0265 (16.9)	.0295 (12.2)	.0093 (5.64)
PSQ	-.0003 (11.8)	-.0004 (13.5)	-.0005 (14.4)	-.0000 (1.07)
XSP	- -	- -	.0001 (1.12)	- -
WTIME	1.4567 (60.0)	1.4079 (54.8)	1.3473 (53.8)	1.8198 (57.2)
GEO1	- -	- -	- -	-.1340 (5.31)
GEO2	- -	- -	- -	-.0449 (2.57)
GEO4	- -	- -	- -	-.1132 (4.65)
GEO5	- -	- -	- -	-.0244 (0.98)
GEO6	- -	- -	- -	-.1520 (6.92)
TYPE	- -	- -	- -	.1047 (7.49)
IND1	- -	- -	- -	-.7910 (24.6)
IND2	- -	- -	- -	.2261 (4.48)
IND3	- -	- -	- -	.0521 (0.61)
IND4	- -	- -	- -	.1505 (3.81)
IND6	- -	- -	- -	.2871 (12.4)
IND7	- -	- -	- -	.0504 (2.40)
IND8	- -	- -	- -	-.2129 (10.9)
IND9	- -	- -	- -	.0165 (0.49)
IND10	- -	- -	- -	-.0255 (1.31)
MAJ	- -	- -	- -	-.1612 (6.76)
IM1	- -	- -	- -	.0070 (0.23)
IM2	- -	- -	- -	-.0282 (1.50)
IM3	- -	- -	- -	.0260 (0.83)
ETH2	- -	- -	- -	.0186 (1.24)
ETH3	- -	- -	- -	.0221 (0.86)
ETH4	- -	- -	- -	-.0264 (0.38)
ETH5	- -	- -	- -	.0472 (0.99)
ETH6	- -	- -	- -	.6521 (8.57)
ETH7	- -	- -	- -	-.0279 (0.91)

[a]Main sample, 22,682 observations

[b]The first figure in each set is a regression coefficient; the second, in parentheses, is the corresponding asymptotic t ratio, written in absolute terms.

the present estimates bear some resemblance to the OLS results of Mincer, though they exceed even the latter by a significant margin.

Comparing the OLS and 3SLS estimates of the hours coefficient suggests that there is indeed a substantial endogeneity bias in the former and that the direction of this bias is negative. Unfortunately, there is no general, *a priori* econometric prediction against which to test the preceding result.

The return to schooling implied by (ME1) is about 1.5 percentage points lower than the corresponding OLS estimate. Proportionately, the experience coefficients shrink by an even greater amount. The one attached to the squared term, which measures the concavity of the experience profile, turns out to be very small indeed. Both results no doubt reflect the increased importance of the hours term and the fact that it depends, in the other equation, upon age and schooling.

The concavity of the experience profile is, of course, a major implication of the human-capital model. Yet, the degree of concavity registered in (ME1), or in any of the structural earnings equations, does not provide especially strong support for the theory. On-the-job investment, if it is indeed the key factor in shaping the experience profile, must not decline very rapidly over the life cycle; but in that case, it must not *begin* at a very high level either, since the model requires that investment cease on or before retirement. Much of the observed concavity in earnings profiles is apparently due to the behaviour of hours.

TABLE 27

SIMULTANEOUS ESTIMATES:[a] HOURS

Right-Hand Variables	Equations[b] (dependent variable = WTIME)			
	(MH1)	(MH2)	(MH3)	(MH4)
Constant	-1.1052 (27.3)	-.7444 (6.43)	-.6357 (6.48)	-.8333 (14.0)
S	-.0158 (13.4)	-.0247 (14.0)	-.0189 (9.97)	-.0180 (12.6)
AGE	.0077 (4.45)	-.0110 (2.53)	-.0179 (4.23)	.0017 (0.67)
ASQ	-.0001 (5.15)	.0001 (2.05)	.0002 (4.04)	-.0000 (2.58)
ZINC	.4209 (11.8)	- -	.8465 (10.8)	.4354 (10.2)
INC	- -	.6379 (9.11)	- -	- -
TMARG	- -	.2494 (0.69)	- -	- -
XINCOTHDI	.0340 (1.05)	-.1040 (1.31)	-.2616 (4.79)	.1441 (5.37)
DI	.4561 (6.88)	.0438 (0.32)	-.2681 (1.95)	.3450 (4.80)
GEO1	-.0260 (3.14)	-.0344 (0.22)	-.0047 (0.39)	.0747 (5.28)
GEO2	-.0061 (0.52)	-.0368 (1.70)	-.0990 (5.14)	.0664 (5.21)
GEO4	-.0138 (1.73)	.0084 (0.63)	.0353 (3.02)	.0576 (4.25)
GEO5	-.0062 (0.83)	.0045 (0.52)	.0236 (2.71)	.0110 (0.85)
GEO6	-.0336 (5.02)	-.0235 (3.15)	-.0135 (1.79)	-.0953 (8.47)
TYPE	.0158 (3.63)	.0015 (0.27)	.0046 (0.80)	-.0511 (6.74)
IND1	.0099 (0.80)	.0730 (3.36)	.1003 (4.51)	.3811 (18.7)
IND2	-.0405 (2.66)	-.0360 (2.15)	-.0178 (1.06)	-.1495 (5.28)
IND3	-.1404 (4.84)	-.0417 (1.07)	.0229 (0.54)	-.1203 (2.56)
IND4	.0319 (2.64)	.0001 (0.00)	.0042 (0.28)	-.0760 (3.61)
IND6	-.0568 (7.21)	-.0476 (4.82)	-.0215 (1.96)	-.1766 (14.3)
IND7	-.0003 (0.05)	-.0073 (1.12)	-.0071 (1.02)	-.0287 (2.63)
IND8	.0099 (1.63)	.0288 (3.54)	.0374 (4.50)	.1150 (10.9)
IND9	.0101 (1.02)	.0150 (1.43)	.0337 (3.01)	.0016 (0.89)
IND10	-.0368 (5.13)	-.0070 (0.63)	.0014 (0.12)	-.0045 (0.38)
MAJ	.0391 (4.49)	.0173 (1.37)	-.0121 (1.00)	.1130 (8.68)
HEAD	.0256 (1.77)	.0290 (1.34)	.0619 (3.81)	-.0176 (1.73)
FAMSIZ	-.0058 (2.69)	.0029 (0.56)	.0113 (3.38)	-.0100 (5.89)
USMAR	.0297 (2.28)	.0213 (1.05)	.0029 (0.19)	-.0018 (0.18)
ETH2	.0015 (1.32)	-.0025 (0.84)	-.0057 (1.73)	-.0113 (1.42)
ETH3	.0004 (0.18)	.0038 (0.78)	.0008 (0.14)	-.0089 (0.65)
ETH4	.0249 (3.94)	.0121 (0.82)	.0040 (0.25)	.0464 (1.26)
ETH5	-.0181 (3.71)	.0006 (0.06)	.0087 (0.74)	-.0512 (1.99)
ETH6	-.0621 (6.21)	.0069 (0.33)	-.0382 (1.76)	-.3332 (8.21)
ETH7	.0130 (4.65)	.0061 (0.99)	.0053 (0.76)	.0205 (1.25)
OC1	-.0420 (4.22)	-.0262 (1.36)	-.0697 (3.27)	-.0963 (7.45)
OC2	-.0304 (4.58)	-.0205 (1.36)	-.0599 (3.96)	-.0700 (7.77)
OC3	-.0205 (3.00)	-.0199 (1.01)	-.0915 (5.22)	-.0622 (5.98)
OC4	-.0126 (1.39)	-.0186 (0.92)	-.0687 (3.03)	-.0580 (4.27)
OC5	-.0042 (1.26)	-.0039 (0.42)	-.0279 (3.06)	-.0193 (3.73)
OC6	-.0255 (4.50)	-.0201 (1.67)	-.0242 (2.16)	-.0582 (8.85)
OC8	-.0294 (6.98)	-.0034 (0.34)	-.0163 (1.54)	-.0203 (3.43)
OC9	-.0065 (1.97)	-.0107 (1.15)	-.0312 (3.56)	-.0271 (5.36)
OC10	-.0176 (5.12)	-.0041 (0.43)	-.0407 (4.48)	-.0251 (4.78)
OC11	-.0106 (3.48)	-.0093 (1.13)	-.0218 (2.64)	-.0203 (4.36)
OC12	-.0164 (5.07)	-.0091 (1.01)	-.0345 (4.13)	-.0296 (6.16)

[a]Main sample, 22,682 observations

[b]The first figure in each set is a regression coefficient; the second, in parentheses, is the corresponding asymptotic t ratio, written in absolute terms

In this connection, it must be understood that the predictions of the structural equations do not relate to the experience profiles one might casually observe and plot. To obtain the counterparts to observation, we must compute the earnings reduced-form equation by substituting (MH1) into (ME1), bearing in mind that AGE = P + S + 5.67, that ASQ = AGE^2, and that ZINC = INC + TMARG. The implied reduced-form coefficients of P and PSQ are 0.0474 and -0.0009 respectively.[16] Those values are only a little smaller than these encountered in the corresponding OLS equation, (CP5)--a fact which indicates rough consistency on the part of the simultaneous estimates.

The reduced-form coefficients suggest that, on average, earnings peak at 27.8 years of experience, or very near the OLS estimate. The structural coefficients place the earnings peak at 35.8 years. For mean-schooled individuals, this point corresponds to 52 years of age. In comparison, hours reach their peak in (MH1) at 30 years of age. This finding is obviously consistent with the prediction of the life-cycle model that the peak in hours comes before the peak in the wage rate. Since hours are declining when earnings peak (that is, at age 52), it follows that the wage rate must still be rising and that it will attain its own peak, if at all, somewhat later.

Actually, since $d \cdot \ln W_a/dp = d \cdot \ln Y/dp - d \cdot \ln h/dp$, one can easily calculate the peak-wage year of experience using the same structural coefficients just employed. Substituting for the two derivatives on the right-hand side, setting the difference equal to zero, and solving for p (the theoretical counterpart of P), one arrives at a figure of 51 years. This point corresponds to age 67 for individuals with mean schooling. In other words, according to the structural estimates of (ME1)-(MH1), wage rates do not reach a peak or decline at all prior to the normal age of retirement. This result agrees, more or less, with Mincer's observation concerning the "weekly earnings" of U.S. males.[17] However, it does not offer much comfort to the human-capital theorist. According to the model, self-investment should not be propelling wages upward when the individual is close to retirement, particularly if depreciation is significant. On the other hand, since the slope of the wage profile is rather slight--one might almost call it flat--in the years approaching retirement, one could still argue on behalf of the theory that investment and depreciation both simply approximate zero during this stage of the life cycle. Such an interpretation, though logically admissible, serves mainly to illustrate how difficult it is to submit the human-capital model to the legitimate jeopardy of scientific falsification.

Focussing on (MH1) alone, we find that the coefficient of ZINC is positive and rather large in absolute terms. On the basis of (66) the implied estimate of δ is 0.73. Such a high value for the elasticity of hours with respect to wages is certainly surprising when one considers the typical results reported in the labour-supply literature. The most common finding for males appears to be that the wage elasticity is negative. The present result therefore raises some suspicion. It must be emphasized, however, that (MH1) makes no pretense at being an identified labour-supply function.

One may think of several reasons to account for the seemingly large value of δ, though none is altogether pleasing. At some level of intuition, it is not surprising that the coefficient of ZINC (and hence δ) is large, since ignoring taxes, we are actually regressing $\ln h$ on the variable $(\ln W + \ln h)$. There would appear to exist a strong tendency for this sum and $\ln h$ to be positively correlated. For many of the labour supply studies, which use wage rates rather than earnings, there is the opposite tendency: h is regressed on Y/h. In both cases, the econometric problem is essentially one of endogeneity. Since the existing studies rely mainly on OLS estimates, bias and inconsistency are to be expected. Here, however, endogeneity receives explicit treatment; thus if the present approach has been successful, inconsistency--and perhaps bias, given the large sample--will have been avoided.

On a more rigourous level, it turns out that in the general case, with several exogenous variables and correlated errors in the structural equations, nothing can be proven about the direction of bias in the coefficient of ZINC. In at least one simplified case, it appears that the direction of bias is indeed positive.[18] A comparison of the OLS estimates in Appendix IV and the present 3SLS results tends to confirm this suggestion. The 3SLS procedure yields a fall in the ZINC, coefficient, though not one of sufficient magnitude to turn δ negative.

Another factor in the present outcome may be the imposition of the constant-elasticity functional form, which has been little used in the existing research. Differences in functional form can obviously have a profound effect upon results. It is not difficult to imagine a labour-supply curve which reverses slopes part way through its range, yielding a positive elasticity estimate with the log-linear specification and a negative elasticity estimate with some other form. A supply curve of this sort, which seems theoretically plausible, may also give contradictory results for different samples or data sets if these are drawn for some reason from different parts of the range.

Finally, it is worth repeating that the hours equation may not be "strongly identified," in the sense that its structure is unquestionably revealed by variables which produce broad and precise shifts in the earnings equation. The possibility exists that in computing the hours regression, we are to a great extent merely running the earnings regression in reverse. A strong positive relationship between wages and hours in the earnings regression would then carry over into the hours estimates. Although this consideration tends to limit interest in the latter, it does not affect the validity of results yielded by the earnings equations.

With regard to the remaining coefficients in (MH1), (66) implies that all must be multiplied by $(1 + \delta)$ to obtain estimates of the structural parameters comprising γ. Even if $(1 + \delta)$ is as large as previously indicated (that is, 1.73), only three of the corrected estimates surpass 0.1 in absolute value.[20] Since the raw coefficients change a good deal in any event as one moves across the table, further calculations are left at this stage to the interested reader.

Before we turn to the other equations, some additional features of (MH1) deserve comment. Note first of all that the raw coefficient of the income-intercept term (XINCOTHDI) is positive but (asymptotically) insignificant--not an uncommon result in the orthodox labour-supply literature. If one were interpreting (MH1) as an identified labour-supply schedule, theory would of course predict a negative coefficient as long as leisure is a normal good.

Schooling, unexpectedly, reduces time worked, both here and in the single-equation estimates displayed in Appendix IV. It would appear that any advantage which the more schooled hold over the less schooled in avoiding unemployment is negated by differences between these groups in labour-supply behaviour or in the time-worked characteristics of their respective jobs. One must be alert, however, to the possibility that schooling, being related directly to earnings, is merely acting as an earnings proxy, thus counterbalancing the latter to some degree and making the functional form less constrained.

As one might casually have forecast, self-employment increases time worked. Though a number of other variables in (MH1) likewise display significance, their coefficients proved generally rather sensitive to the particular specification in force and are therefore best considered in light of all the results.

Further Experiments

Equations (ME2)-(MH2) show the effects of inserting INC and TMARG separately in the hours regression. On the earnings side, the coefficients change very little and, hence, require no additional comment. However, in the hours regression itself, the modification is crucial. The coefficients of INC and TMARG, first of all, are significantly different from each other, contrary to standard theoretical reasoning. Taxes appear much less important than gross earnings. Nevertheless, in view of the problems in estimating the tax rate and in purging TMARG of its endogeneity, one cannot treat this result as more than suggestive.[21] Second, in response to the change, the coefficients of AGE and ASQ switch signs, indicating a convex rather than a concave structural profile of hours. Third, most of the other coefficients become asymptotically less significant than in (MH1).[22] The use of the *two* income-related terms in place of ZINC tends, it seems, to overpower the other variables.

Equations (ME3)-(MH3) restore the use of ZINC in order to investigate the effects of SSQ and XSP in the earnings regression. As before, the coefficient of SSQ is significantly positive, but that of XSP is insignificant. For individuals with mean levels of schooling and experience, the implied rate of return to the former is 6.2%--again, somewhat lower than estimated by OLS. This figure rises (falls) by 0.5 percentage points for each year of schooling above (below) the mean. The reduced-form earnings profile turns out to be convex rather than concave,[23] thereby casting general doubt upon this version of the model. As in (MH3), the structural profile of hours is also convex.

We come now to the expanded earnings function, (ME4). The insertion here of twenty-five additional variables causes some marked changes in the coefficients upon which we have been focussing. The indicated return to schooling falls by approximately one further percentage point to 5.3%. The increases in earnings on account of experience become very small indeed, and the concavity of the earnings profile (as registered in the structural estimates) disappears. As a compensation, the importance of hours worked greatly increases. The elasticity of wages with respect to hours is given as 1.82. Overall, then, the influence attributed to the orthodox human-capital proxies, S, P, and PSQ, when these change ceteris paribus, is substantially diminished. Though it is arguable, because of linked mobility patterns, whether ceteris-paribus measurements are actually legitimate,[24] the present estimates serve to show the effect of not conceding to the human-capital variables, as Mincer and others do, the "benefit of the doubt."

It will be observed that, among the variables added in (ME4) to the basic human-capital specification, the coefficients of many remain very sizable. For example, residence in Atlantic Canada (GEO1) is a disadvantage worth 2.6 years of schooling; residence in British Columbia (GEO6) is an advantage worth 2.9 years. Employment in agriculture (IND1) is an immense handicap (79% of reference-group earnings), whereas employment in construction (IND6) yeilds top earnings (29% more than in manufacturing). Period of immigration (IM) is not significant, but rural or small-town residence (TYPE) and self-employment (MAJ) continue, as in the OLS results, to exact substantial earnings penalties.

The coefficients of ethnic group (ETH) perhaps deserve special comment. The one pertaining to individuals of Jewish descent (ETH5) remains positive but is no longer significant, as it was in the OLS regressions. The coefficient pertaining to Native Indians (ETH6) is the only one which is significant here, and it is both positive and very large, contrary no doubt to one's casual predictions. It must be remembered, however, that the coefficient in question measures the effect of Native Indian origin with other variables such as schooling, experience, hours, location, and industry *held constant*--a situation we do not casually observe in the real world. The calculated reduced-form coefficient is much smaller (0.0457), since hours at least are permitted to vary; still, for the most part, ceteris paribus applies. Though the present result may yet seem anomalous, it receives some support from the findings of Haessel and Kuch.[25] One might speculate that, as an apparently disadvantaged group, Native Indians benefit particularly from socially or institutionally standard rates of pay, which they receive when employed, despite inferior qualifications.

As for the hours structural equation, (MH4), it will be observed that in every case but one, the signs of the added variables are the reverse of those in the earnings structural equation. Within particular categories, hours worked tend to offset high earnings. This result may be a further clue to the apparent high value obtained for the coefficient of ZINC. When hours are low and earnings high, implicit or actual wage rates per hour must be high as well. We thus come upon some indication of a negative relationship between wage rates and hours. If negative aspects of the overall relationship are closely linked with the added variables (GEO, TYPE, IND, et cetera), these will tend to reflect the negative side, leaving the coefficient of ZINC relatively large. This tendency will operate to some extent even when the variables in question do not appear in the earnings equation; then, since fewer attributes are held constant across the entire system, and the need for offsetting coefficients is less pronounced, one would expect those which remain in the hours equation to lie closer to zero. This pattern does emerge in the comparison of (MH4) and (MH1). However, the change in the coefficient of ZINC, while in the anticipated direction, is rather small. One can say only that adding variables to the system--holding their influence constant, in other words--may be in part responsible for the finding with respect to ZINC.

The observation that wages and hours are broadly offsetting when viewed across regions and industries tends to redeem the speculation concerning seasonality made earlier in connection with the OLS estimates. If seasonality is indeed the ruling factor in the creation of offsetting wage differentials, it is by no means surprising that we should observe the effect through regions and industries, which seasonality strikes unevenly.[26] In the OLS equations the seasonal effects cannot manifest themselves except through the coefficient of WTIME. In the 3SLS equations the latter is free to reflect other links between wages and hours, such as the rates earned moonlighting, the premium for overtime, and the unmeasured ability variables which influence wages and hours in common.

It is worth noting, finally, that (MH4), like all the other structural equations, displays scant concavity in the implied experience or age profile. There is a very flat peak in hours at 10.0 years of experience. This result nevertheless satisfies the prediction of the life-cycle model, since earnings peak (structurally) well beyond the relevant range--at 141 years, to be precise. From the standpoint of the computed reduced form, hours and earnings peak at 15.8 and 20.4 years of experience respectively. These points come a little earlier than calculated previously. One may wonder, given that the change in specification has been to hold additional variables constant, whether individuals thus use geographic and interindustrial mobility to stave off earnings and hours peaks. If such moves benefit individuals at various points in their life, one should indeed notice a hastening of the peaks when this recourse is disallowed statistically in cross-section.

APPENDIX V

ESTIMATES OBTAINED BY ITERATIVE THREE-STAGE LEAST SQUARES[1]

$$\text{(ME5)}\quad \text{INC} = \underset{(46.3)}{1.1920} + \underset{(33.5)}{0.0589}\ \text{S} + \underset{(10.3)}{0.0087}\ \text{P} - \underset{(5.28)}{0.0001}\ \text{PSQ} + \underset{(66.8)}{1.5019}\ \text{WTIME}$$

$$\begin{aligned}
\text{(MH5)}\quad \text{WTIME} = {} & \underset{(20.4)}{-0.6742} - \underset{(29.1)}{0.0357}\ \text{S} - \underset{(4.68)}{0.0062}\ \text{AGE} + \underset{(1.10)}{0.0000}\ \text{ASQ} \\
& + \underset{(27.1)}{0.6324}\ \text{ZINC} + \underset{(10.4)}{0.1451}\ \text{XINCOTHDI} + \underset{(6.66)}{0.2480}\ \text{DI} \\
& + \underset{(2.19)}{0.0081}\ \text{GEO1} + \underset{(7.00)}{0.0333}\ \text{GEO2} + \underset{(1.60)}{0.0056}\ \text{GEO4} + \underset{(0.72)}{0.0017}\ \text{GEO5} \\
& - \underset{(3.46)}{0.0078}\ \text{GEO6} + \underset{(1.16)}{0.0020}\ \text{TYPE} + \underset{(0.28)}{0.0024}\ \text{IND1} - \underset{(1.61)}{0.0076}\ \text{IND2} \\
& - \underset{(3.42)}{0.0404}\ \text{IND3} + \underset{(0.59)}{0.0027}\ \text{IND4} - \underset{(3.03)}{0.0098}\ \text{IND6} + \underset{(1.25)}{0.0025}\ \text{IND7} \\
& - \underset{(1.62)}{0.0043}\ \text{IND8} - \underset{(2.54)}{0.0079}\ \text{IND9} - \underset{(1.98)}{0.0071}\ \text{IND10} + \underset{(6.22)}{0.0195}\ \text{MAJ} \\
& - \underset{(5.73)}{0.0258}\ \text{HEAD} - \underset{(11.6)}{0.0096}\ \text{FAMSIZ} - \underset{(4.56)}{0.0198}\ \text{USMAR} \\
& - \underset{(2.03)}{0.0023}\ \text{ETH2} + \underset{(0.27)}{0.0005}\ \text{ETH3} + \underset{(0.50)}{0.0027}\ \text{ETH4} + \underset{(5.13)}{0.0200}\ \text{ETH5} \\
& - \underset{(4.05)}{0.0361}\ \text{ETH6} - \underset{(0.32)}{0.0007}\ \text{ETH7} + \underset{(5.26)}{0.0368}\ \text{OC1} + \underset{(3.19)}{0.0160}\ \text{OC2} \\
& + \underset{(2.38)}{0.0141}\ \text{OC3} + \underset{(6.43)}{0.0475}\ \text{OC4} + \underset{(0.53)}{0.0016}\ \text{OC5} + \underset{(1.70)}{0.0063}\ \text{OC6} \\
& - \underset{(2.27)}{0.0086}\ \text{OC8} + \underset{(0.58)}{0.0017}\ \text{OC9} - \underset{(0.16)}{0.0005}\ \text{OC10} - \underset{(0.33)}{0.0009}\ \text{OC11} \\
& + \underset{(0.03)}{0.0001}\ \text{OC12}
\end{aligned}$$

Number of observations = 22,682

Number of iterations = 11

---

[1] Figures in parentheses are asymptotic t ratios, written in absolute terms.

NOTES

CHAPTER V

[1]The symbols Y and h now stand for vectors, both of then $N \times 1$, N being the number of observations in the sample. The previously defined vectors X_i' and Z_i', $i = 1, 2, \cdots, N$ (i formerly suppressed) make up the rows of X and Z respectively.

[2]As a general reference the reader may wish to consult J. Johnston, Econometric Methods (second edition; New York: McGraw-Hill Book Co., 1972), pp. 395-398.

[3]For proof see Phoebus J. Dhrymes, "Small-Sample and Asymptotic Relations between Maximum-Likelihood and Three-Stage Least-Squares Estimators," Econometrica, XLI (March, 1973), pp. 357-364.

[4]See Albert Madansky, "On the Efficiency of Three-Stage Least-Squares Estimation," Econometrica, XXXII (January-April, 1964), 51-56.

[5]Further discussion on this point is provided by Terence J. Wales and Alan D. Woodland, "Labour Supply and Progressive Taxes," Review of Economic Studies, XLVI (January, 1979), 83-95. Besides dealing with endogeneity, these authors investigate what they call "specification error," which results from a stochastic discrepancy between the actual and desired labour supply of the individual. However, this problem really arises only within an explicit utility framework, when one is assuming the identification of a labour-supply function.

[6]See Wales, "Estimation of a Labour-Supply Curve for Self-Employed Business Proprietors."

[7]Since ZINC stands for $\ln(1-\tau)\,Y = \ln(1-\tau) + \ln Y$ the determinants of $\ln(1-\tau)$ and $\ln Y$ at least *combine* additively in the present formulation. Note from the definitions of Chapter III that ZINC is furthermore an additive component in the calculation of XINCOTHDI.

[8]The list reads as follows: S, SSQ, P, PSQ, XSP, LEN, GEO, TYPE, IND, MAJ, OC, HEAD, FAMSIZ, USMAR.

[9]The reader may verify the point by consulting Figure 2 along with the definitions of Chapter III.

[10]See again Wales, op. cit.

[11]Seven hours categories and five weeks categories yields thirty-five possible combinations.

[12]The interval is half closed because individuals who worked zero hours were previously excluded from the sample. Such exclusion gives rise to a problem now known in the literature as "sample selectivity bias." See Reuben Gronau, "Wage Comparisons - A Selectivity Bias," Journal of Political Economy, LXXXII (November/December, 1974), 1119-1143. Estimates are "biased" in the sense that they are *conditional* upen the individual's working at some time during the measurement period; hence, they may not hold in the aggregate and maybe misleading for policy purposes if not correctly interpreted. See also Michael J. Boskin, "The Economics of Labor Supply," in Income Maintenance and Labor Supply, edited by Glen G. Cain and Harold W. Watts (Chicago: Rand McNally College Publishing Company, 1973).

[13]For a more sophisticated treatment see Giora Hanoch, A Multivariate Model of Labor Supply: Methodology for Estimation (Santa Monica: The Rand Corporation, September, 1976).

[14]See Arnold Zellner and Henri Theil, "Three-Stage Least Squares: Simultaneous Estimation of Simultaneous Equations," Econometrica, XXX (January, 1962), 54-78.

[15]Cf. Equation (CP5), Table 6.

[16]Note that the entries in the tables have been rounded. Hence, the results stated here and below may not appear entirely consistent with the reported figures.

[17]Schooling, Experience, and Earnings, p. 70.

[18]Consider the highly abbreviated model

$$\ln Y = x + (1+\theta)\ln h + \tilde{u}_1 ,$$

$$\ln h = a + D \ln Y + \tilde{u}_2 ,$$

where $D \equiv \frac{\delta}{1+\delta}$, a is a constant, and x represents the sum of all factors determining $\ln \hat{W}_a$. Assume that $\tilde{u}_1$ is uncorrelated with both x and $\tilde{u}_2$ and that the latter are themselves uncorrelated. This case is just slightly more general than one treated by Johnston, op. cit., pp. 342-344. The reduced forms are:

$$\ln Y = \frac{1}{1-D(1+\theta)} \; [a(1+\theta) + x + \tilde{u}_1 + \tilde{u}_2(1+\theta)]$$

$$\ln h = \frac{1}{1-D(1+\theta)} \; [a + Dx + D\tilde{u}_1 + \tilde{u}_2] \quad .$$

Following Johnston, we may compute moments (denoted m_{ij}), and probability limits. In light of our assumptions, we find that the relevant moments are

$$m_{yy} = \frac{1}{[1-D(1+\theta)]^2} [m_{xx} + m_{\tilde{u}_1\tilde{u}_1} + m_{\tilde{u}_2\tilde{u}_2}(1+\theta)^2]$$

$$m_{\ h} = \frac{1}{[1-D(1+\theta)]^2} [Dm_{xx} + Dm_{\tilde{u}_1\tilde{u}_1} + m_{\tilde{u}_2\tilde{u}_2}(1+\theta)] \quad .$$

The OLS estimate of D is $D = m_{yh}/m_{yy}$. It follows that

$$\text{plim } \hat{D} = \frac{D\bar{m}_{xx} + D\bar{m}_{\tilde{u}_1\tilde{u}_1} + \bar{m}_{\tilde{u}_2\tilde{u}_2}(1+\theta)}{\bar{m}_{xx} + \bar{m}_{\tilde{u}_1\tilde{u}_1} + \bar{m}_{\tilde{u}_2\tilde{u}_2}(1+\theta)^2}$$

if $\text{plim } m_{ij} = \bar{m}_{ij} < \infty$ for all i, j.

The asymptotic bias is therefore

$$\text{plim } \hat{D} - D = \frac{(1-D)(1+\theta)\,\bar{m}_{\tilde{u}_2\tilde{u}_2}}{\bar{m}_{xx} + \bar{m}_{\tilde{u}_1\tilde{u}_1} + \bar{m}_{\tilde{u}_2\tilde{u}_2}(1+\theta)^2}$$

which is positive as long as $\theta < -1$, since $D = \delta/(1+\delta)$ cannot exceed unity. The reader may verify that if any of the noncorrelation assumptions are violated or if the hours equation contains additional nonorthoganal independent variables, the direction of bias in indeterminate.

[19]See Yoram Barzel and Richard J. McDonald, "Assets, Subsistence, and the Supply Curve of Labor," American Economic Review, LXIII (September, 1973), 621-633.

[20]These are associated with DI (really, just an adjunct to the constant), with IND3 (fishing and trapping, a seasonal industry), and with ETH6 (Native Indians).

[21]In particular, it may arise because of collinearity between TMARG and INC.

[22]Multicollinearity is again a problem in the case of the family-status variables HEAD, FAMSIZ, and USMAR, which contribute initially to the estimation of TMARG.

[23]The reduced-form coefficients of P and PSQ are -0.1146 and 0.0010, respectively.

[24]See again the discussion in Chapter II.

[25]"Size Distribution of Earnings in Canada." In this study the largest coefficients, which are very nearly identical, belong to Native Indians and to Chinese and Japanese. Perhaps on account of small numbers, those pertaining to Native Indians (an intercept dummy and a schooling interaction) are insignificant. Though an exact comparison is impossible because of differences in definition, the ceteris-paribus earnings advantage estimated by the authors for mean-schooled individuals appears to be about the same as that implied by the present structural equations.

[26]Other factors besides seasonality--proneness to strikes, for example--might also fit this criterion; but the alternatives, which may contribute *something* to the explanation, do not account as plausibly as seasonality for the discrepancy between Canadian and American results under OLS.

CHAPTER VI

SUMMARY AND CONCLUSIONS

This chapter provides a condensation of the arguments and inferences stated in the preceding text. It reviews the assorted methodological, theoretical, and econometric objections raised against the human-capital model and attempts, in light of these objections, to place the empirical exercises of the current study in the proper perspective. Results are summarized for a large cross-section of Canadian males who worked in 1970.

CHAPTER I

Three models are considered, in ascending order of their generality: (1) the basic schooling model and (2) the model of postschool investment, both employed by Mincer, and (3) the earnings maximization model, suggested by Ben Porath. The first two deal with investment in human capital at particular stages of the life cycle; the third contains the others as special cases.

The schooling model asserts that proportionate differences in earnings accompany absolute differences in years of formal schooling; that is, $\ln W_s = \ln W_0 + r^e s$, where the parameter governing the relationship,

r^e, is interpreted as the rate of return to education. The assumptions needed to sustain this interpretation are, however, exceedingly powerful. The fundamental postulate is that individuals receive the same capitalized sum in lifetime earnings no matter what their level of schooling.

If the supposed equality of present values is merely conceptual, then r^e is at best an ex post internal rate of return. One must say "at best" because the present-value accounting prescribed by the model is very rough. Schooling is presumed to entail no direct expenditures or subsidies, no present or future nonpecuniary benefits or costs, and no opportunities for part-time employment. Hours of work and the risks of unemployment are held constant, over the life cycle of the individual and across schooling groups. Though estimates based on these assumptions may, even so, provide some useful description of the earnings structure, they cannot be regarded as furnishing *tests* of any maintained hypothesis.

On the other hand, if the equality of present values is presumed to be *actual* , then (subject to the preceding approximations) r^e may be thought of as an ex ante, long-run equilibrium rate of return. Mincer, and other writers of the human-capital school, are nevertheless mainly silent on how the labour market might function to bring about long-run equilibrium. No analysis of individual choice is ever provided within the context of the schooling model, though it is possible to devise one if individuals are assumed to ignore leisure in favour of maximizing a single objective, discounted lifetime earnings. Again, however, the exercise fails to place any important restrictions on the data.

Whereas the supply side of the labour market gains at least a shadowy presence, the demand side suffers complete neglect. Although the schooling model is unquestionably a reduced-form relationship from the labour-market standpoint, no exogenous demand variables appear in it. Market imperfections, associated perhaps with region or industry, are deemed unimportant, as are any quantity imbalances which might cause a "temporary" departure from long-run equilibrium. Since one cannot tell whether long-run equilibrium actually obtains at any given moment of observation, there is no conceivable way of testing the schooling model. That its parameter r^e might supply an adequate ex post empirical description thus remains the strongest admissible claim.

Becker's model of the individual's market for human capital also turns out to be barren of testable implications. It is consistent with any sort of cross-sectional relationship between r^e and the level of schooling. The "interactions model," which Haessel and Kuch derive from it, nevertheless holds some promise. In this analysis, r^e is at least made to depend upon some measurable attributes of the individual. Though the hypotheses linking r^e and these attributes remain essentially ad hoc, they lead one, as theory should, to investigate new dimensions of the empirical earnings structure.

Distributional arguments flowing from the schooling model typically rest on the assumption of independence between schooling and its rate of return. The evidence against such an assumption is, however, very widespread. The only unambiguous implication is that earnings will follow the lognormal distribution (or weaker, be skewed to the right) if schooling follows the normal (is not heavily skewed to the left). Unfortunately, as Oulton points out, there is no theory to specify the distribution of schooling.

The postschool investment model elaborates upon the schooling model by allowing individuals to divide their time between training and pure work in accordance with a second parameter k. Leisure is again held constant, and the model is derived through a series of identities and approximations. It is shown that if k, the propensity to invest in human capital, declines over the life cycle, then the model is consistent with the principal stylized fact concerning age-earnings profiles, namely, that they are concave from below. However, concavity may also be due to biological aging or to costless learning by doing. Only an appeal to competitive equilibrium will rule out the latter. Hence, all the criticisms directed at the schooling model still apply. Empirical "tests" do not discriminate among all three competing hypotheses.

Using the concept of "overtaking," Mincer derives the prediction that the cross-sectional variance of earnings will display a minimum at roughly $p = 1/r^e$ years of experience. This hypothesis is not strongly confirmed by Mincer's data; but since it is conditional upon there being only a small correlation between earnings at school leaving and the propensity to engage in postschool investment, the model proves in the end to be immune from falsification on this account.

The Ben Porath model seeks to provide a behavioural theory of k based on a formal analysis of the conditions for maximizing discounted lifetime earnings. It is shown that the optimal values of k do in fact decline over the life cycle, the reason being that the time period over which to amortize successive investments becomes increasingly short. Therefore, present-value maximization is generally consistent with the concavity of age-earnings profiles. Yet, in detailed testing, the shapes of these profiles do not conform to expectations. It has been suggested that the fault lies in the "neutrality hypothesis," which restricts the form of the human-capital input in its alternative uses. However, without the neutrality hypothesis, the model is untestable.

Although the three models surveyed may be useful as an aid to thought and as a framework for empirical description, they fail, for the most part, to generate critical hypotheses by means of which to test the central notion that earnings are the result of individual investment decisions.

CHAPTER II

Even if implementation of the various models turns out to be merely an exercise in description, it is still necessary to consider the problems which may hinder unbiased estimation. Descriptive results, even if *correctly* interpreted as such, should not be misleading from a quantitative point of view.

By means of simple regression, Mincer estimates the return to schooling at 7%--a figure much below the values obtained directly from age-earnings profiles in other studies. He attributes the apparent downward bias to the omission of experience (postschool investment), which is negatively correlated with schooling. It is argued here that any net bias may involve several factors: (1) individual variation in rate of return (schooling coefficient), (2) the endogeneity of schooling, (3) expectations and growth, (4) omission of ability and family background, (5) omission of other variables.

If the individual rate of return falls as schooling increases, the simple-regression estimate of r^e will have a downward bias. Yet, in the case of Canada, one might look for an upward bias, since existing research gives some hint of rising returns. Mincer argues, with respect to the United States, that the apparent fall in the rate of return is due to the variation in weeks worked. It may not be ligitimate, however, to estimate r^e with weeks worked held constant. An alternative approach is to account explicitly for individual differences in r^e, either by letting the variable "years of schooling squared" appear in the regression, or by resorting to the more elaborate interactions framework. In general, the power of the human-capital model suffers to the extent that r^e turns out not to be a stable parameter.

If schooling is really an endogenous variable, the estimated return will again be subject to bias. Proponants of the model must therefore confront a dilemma: endogenous schooling leads to biased estimation, but exogenous schooling means that there is no market mechanism to enforce long-run equilibrium.

Expectations, mainly with regard to the growth of wages, must also be considered. If long-run equilibrium is assumed, the schooling coefficient will measure only the difference between the net rate of return and the average expected rate of real growth (that is, $r^e - g^*$). An underestimate of the former, caused by misinterpretation, may thus occur. Age and place of highest grade might serve as proxies for the state of expectations at a particular time in a particular locale.

Among all the potential sources of bias in estimating the rate of return, the one which has received the most attention has been the omission of ability and family background. It is argued that if ability and family background have an independent effect on earnings, and if these variables are positively correlated with schooling, then their omission will bias the schooling coefficient upward, as the latter "picks up" earnings variance which is not causally attributable to it. The census data used here and in the comparable study by Mincer do not, of course, provide the ability and background variables with which to investigate this problem further.

It *is* possible to investigate potential biases from the omission of other variables. It is argued here that industry, occupation, and place of residence may capture components in the apparent rate of return which are the result of market imperfection, short-run disequilibrium, and previously ignored nonpecuniary factors. Such components will not be available to every investor in schooling. Though it may be that schooling is a *prior* cause of industrial, occupational, and geographic mobility, one cannot assume that the variables mentioned have *no* independent effect. The human-capital model makes this assumption and thus attributes *all* the doubtful earnings variance to schooling.

Mincer holds weeks worked constant, but none of the suggested variables. When he inserts weeks worked, the implied rate of return to schooling falls. It turns out that the elasticity of earnings with respect to weeks is greater than unity.

Implementation of the postschool investment model requires, first, that one estimate the amount of time an individual has spent on the job (his "experience") and, second, that one specify the proportion of time (k) devoted in each period to training. To estimate experience, Mincer and others use age minus schooling minus five. This proxy assumes no unemployment or nonparticipation in the labour force, together with constant hours. The associated errors of measurement may bias the schooling coefficient up or down in the eventual formulation; however, empirical evidence suggests an upward bias. To specify the time profile of k, Mincer proposes two functions, one of which declines linearly, and the other, exponentially. The former leads to a quadratic estimating equation; the latter, to another exponential. Neither specification quite matches the theoretical form implied by the Ben Porath model, although both may give a tolerable approximation. The exponential form has the advantage of identifying all the parameters of the empirical model.

Besides holding experience constant in the preceding parametric fashion, Mincer uses the alternative method of applying the schooling model to a single experience cohort, the one estimated to be at overtaking. Within this cohort, earnings differentials are thought to be entirely attributable to schooling. In either case, the schooling coefficient rises considerably as predicted.

Implementation of the Ben Porath or income-maximization model is complicated by the unavoidable nonlinearity of functional forms. This problem dictates relatively small sample sizes with few variables. As a result, it has been impossible to test hypotheses of real interest--those which link the theoretical parameters to individual attributes. Attempts at implementation have been, to a great extent, exercises in curve fitting, as earnings are regressed on age or experience transformed in diverse ways. "Reasonable" parameter estimates are then taken as confirmation of the theory.

CHAPTER III

Here, the previously surveyed aspects of Mincer's empirical work on the human-capital model are reproduced using Canadian microdata. The standard earnings function is then expanded by means of additional variables, the aim being, on the one hand, to provide an improved description of the Canadian labour economy and, on the other, to establish an alternative benchmark against which to judge the orthodox specification.

The principal data source for this effort was the one-in-one-hundred Public Use Sample drawn from the 1971 census. The working sample comprised 22,682 out-of-school males who were employed at some time during 1970 in any of the 10 identifiable industries making up the private sector. Each observation consisted of individual data on 168 variables.

Results for the basic schooling model were virtually identical to those reported by Mincer. The schooling coefficient or "rate of return" was measured at 7%. The simple regression explained 7% of the earnings variance or "inequality."

As in Mincer's work, experience was held constant in three ways: by examining the overtaking cohort and by estimating, first, the exponential, and then, the quadratic specification. The overtaking subsample consisted of 1,238 individuals with 7-9 years of experience. For this group the schooling coefficient reached 10% but fell by one-quarter when weeks worked were held constant. The insignificance of schooling squared implied, according to the orthodox interpretation, that the return to schooling did not vary. However, the level of the return was not entirely consistent with the definition of the overtaking set laid down in part with the aid of Mincer's reciprocal rule of thumb. The elasticity of earnings with respect to weeks was not significantly different from unity.

The exponential form of the experience profile was investigated by iterating a linear equation for different values of β, the exponential rate at which k declines over the life cycle. Since we must have

$0 \leqq k \leqq 1$ together with a positive return on postschool investment ($r^x > 0$), it was possible to deduce certain reasonableness restrictions with which to screen the estimates. In one variant of the model reasonable coefficients were implied for β in the 0.15-0.20 range, but the results proved far too unstable to use in computing estimates of r^x and k_0 (the initial propensity to invest). In another variant β would have had to be somewhat less than 0.05. As for the other coefficients, that of schooling squared was significantly positive; that of weeks was significantly less than unity.

When the quadratic functional form was used to portray experience, the implied rate of return to schooling was 8.7%. With weeks held constant, the figure declined by about one-sixth to 7.2%. The coefficient of schooling squared, though relatively small, was significantly positive whether or not the weeks variable was included. Earnings peaked at 29-30 years of experience--a little earlier than in the United States sample. Experience profiles had a slight tendency to converge over the life cycle, just as Mincer observed. Each additional year of schooling postponed the earnings peak by only 0.6-0.7 years. Mincer's assumptions with respect to depreciation and the length of the net investment stage produced estimates of 7.7% for r^x and 0.54 for k_0'. However, a wide range of values were obtained by varying these assumptions within reasonable limits.

Generally speaking, the introduction of experience, by whatever means, had considerably less effect here in raising the coefficient of schooling than it did in Mincer's research. On average, rates of return (if one chooses to interpret the schooling coefficients as such) appear to be lower in Canada than in the United States. As noted, however, differences in sample composition and in time period may contribute to this result. There is nevertheless a firm contrast in the tendency of Canadian returns to rise with the level of schooling and in the observation that, over the full sample, earnings did not rise in proportion to the number of weeks worked.

Over both the expanded and the orthodox earnings functions, implied rates of return varied from 6.9% to 8.91% with hours ignored and from 6.03% to 7.75% with hours held constant. Corrected for anticipated real growth, these values exceeded the most recent estimates of Statistics Canada but were still well short of those computed by Podoluk a decade earlier. Additional variables in the expanded regressions did not account for rising returns until occupation was introduced; then, schooling squared became insignificant. Returns that rise in cross section are not, of course, inconsistent with a competitive equilibrium. The elasticity of earnings with respect to hours was considerably less than unity in all the single-equation estimates.

Among the added variables, "long" vocational training was associated with an earnings premium of 8% to 18%, depending on the specification. Industry and place of residence, the variables taken here to represent market imperfections, disequilibria, and nonpecuniary factors, were highly significant, contributing almost as much to R^2 upon addition as schooling, and somewhat more upon deletion from the full model. In fact, the deletion of all human-capital variables lowered R^2 by 0.042, whereas the deletion of all "unorthodox" variables lowered it by 0.105. This result leaves open to question whether the emphasis accorded human capital in the existing literature has been wholly justified.

A number of variables contributed in only a minor way to the value of R^2 but were nevertheless of some interest. For example, self-employment proved to be a significant earning handicap on balance, as did recent immigration. However, immigrants suffered no lasting disadvantage. Married heads of families turned out to receive 30-31% more than the reference group. Unilingual francophones earned 11-12% less. Among ethnic groups, those of Jewish origin led the ranking. Native Indians fared worst, though not on account of wages that were low (ceteris paribus), but on account of meagre employment.

Finally, a version of the interactions model was estimated to discover whether industry or place of residence affected the earnings potency of schooling and experience. The schooling interactions were not significant; the experience interactions, moderately so. Neither set added impressively to the value of R^2.

CHAPTER IV

The preceding analysis abstracts completely from all planned variation in hours of work. At best, the models deal with the maximization of lifetime earnings rather than with the maximization of utility. However, since time is presumably both an argument of the utility function and an input in the production of human capital, decisions concerning its allocation among work, leisure, and investment are best treated simultaneously.

Three control-theoretic studies of simultaneous decision-making were surveyed in order to compare their assumptions and to obtain predictions with respect to several broad inquiries. It was found that these utility-based analyses tended to undermine the assertion of the simpler human-capital models that investment declines monotonically over the life-cycle. Cases were uncovered in which investment might increase during a given period. The concavity of the earnings profile was therefore seen as depending more heavily than in the earlier models upon the concavity of the hours profile. The latter was forecast to be unambiguously concave. It was deduced in one study that if the market rate of interest exceeded the rate of time preference, then there would be a peak in hours prior to successive peaks in earnings and in wages. Unfortunately, the studies surveyed produced no equations which were amenable to direct estimation, and there was again scant discussion of hypotheses which might associate unobservable parameters with the observable attributes of individuals.

Accordingly, a simplified empirical model of wages and hours was postulated in an attempt to deal with the gross facts involving simultaneity. The practical aim of this two-equation linear model was to obtain estimates of the earnings function which would be free of the bias suspected on account of an endogenous hours variable.

In the proposed specification, earnings were (identically) the product of hours and the average wage. Owing to such things as moonlighting, overtime, and seasonality, the average wage was allowed to depend (stochastically) not only upon schooling, experience, and so forth, but also upon hours worked. The latter was made a (stochastic) function of certain exogenous variables and the marginal wage. The average and marginal wage rates differed both through the dependence of the former on hours and through personal income taxes. Though it could be argued from a life-cycle perspective that, since wage rates are planned, there is no need to include them in the hours equation alongside the age variable, the marginal wage *was* introduced separately in order to represent unforeseen influences, initial endowments, and various unmeasured qualities of the individual. Although the hours equation resembled the "labour-supply" functions frequently estimated in the literature, no attempt was made to press this interpretation in view of the strong assumptions required to guarantee the identification of a pure supply relationship.

CHAPTER V

The two-equation model was estimated by the method of three-stage least squares. This procedure allows not only for endogenous variables on the right-hand side but also for the possibility that the error terms may be correlated across equations. Under stated assumptions, the resulting estimates are consistent and asymptotically efficient, though they may differ numerically from those of maximum likelihood.

Further econometric difficulties involved the endogeneity of the tax rate, the limited-dependent-variable status of the hours term, and the identification of the hours equation. Instrumental-variable estimates were used in an attempt to rid the principal tax-related terms of their endogenous components. In view of the obstacles one meets in trying to approximate the progressive tax structure, this effort must unfortunately be judged somewhat speculative. It was pointed out that, strictly speaking, the hours term constituted a discrete and limited dependent variable, but that the problem was less severe than if weeks alone had been employed. A simple but important caveat was entered--namely, that all estimates must be regarded as conditional on participants working positive hours. It was finally noted tht the identification of the hours equation might be somewhat tenuous, since it was obtained through the omission of only two variables, schooling squared and an age-schooling interaction, which were probably of minor importance.

Among the initial findings, the most striking was the rise in the estimated elasticity of earnings with respect to hours--from 0.6-0.7 in the single-equation regressions to 1.4-1.8 in the simultaneous results.

The implied return to schooling fell to 6.3% in the orthodox earnings function, and the experience coefficients were also diminished in size. The structural earnings profile registered very slight concavity, thus casting some doubt on the human-capital interpretation of wage rates. Hours were found to peak before earnings, as the choice-theoretic model suggested; wage rates appeared not to peak in the relevant range, as Mincer discovered in the United States.

The hours equation proved rather sensitive to changes in model specification, yielding in some cases a concave, and in others a convex, hours profile. Moreover, the implied elasticity of hours with respect to the wage rate was a good deal larger than one might have forecast on the basis of conventional labour-supply studies. Problems of estimation bias, differences in functional form, and differences in variables and methods may explain this apparent discrepancy.

Contrary to United States experience, the coefficient of schooling squared was again significantly positive: each additional year of schooling raised the estimated return by 0.5 percentage points. This finding is consistent with the suggestion, occasionally voiced in Canada, that this country has a relative scarcity of workers at the higher levels of education.

The expanded earnings function, estimated simultaneously with hours, provided considerable detail on the pattern of rewards prevailing across the Canadian work force. Two general observations were: that high wages tended to offset low hours, perhaps because of market equalization between jobs with high and low risks of unemployment or high and low seasonality; and that earnings peaks were hastened when geographic and interindustrial mobility was, in effect, disallowed. The estimated mean rate of return to schooling was a mere 5.3%, and the experience profile of earnings was virtually linear rather than strictly concave.

FINAL REMARKS

It has been argued in this study that the human-capital approach to earnings determination lacks the hard testability required of a scientific theory and that it may serve, at best, as a framework for empirical description. A descriptive profile, based loosely on the human-capital paradigm, was therefore drawn to portray the Canadian earnings structure, which has not been analysed extensively in this fashion.

From a purely empirical standpoint, it did not turn out that the orthodox human-capital variables were of *overwhelming* importance. Industry, place of residence, and other factors were also significant; and it cannot be assumed a priori that all are simply means through which individual investment plans are realized, as some have contended. Even if the preceding assertion is correct, one would have to concede on the basis of the present results that mobility with respect to the factors just named is an important concern. If education and mobility are *both* essential for the realization of a given earnings increment, policy initiatives, if any are needed, cannot afford to slight either one.

Private, pecuniary rates of return to schooling, estimated roughly, with hours constant, by the method of semi-log regression, fell in the 5-8% range. This is well below Mincer's estimates for the United States, though it is worth repeating here that the years of observation differed by a decade. If one were to apply Becker's efficiency criterion and thus compare real market rates of return on human and nonhuman capital, one's conclusion would have to be that there is some prima facie evidence of underinvestment in education on the part of individuals. Yet, it is difficult to say what risk premium has been attached to educational investment, and it must be emphasized, lest the reader attempt to make policy inferences, that private and social returns may differ.

Though the present study has succeeded in adducing a number of interesting facts with regard to the Canadian earnings structure, problems remain which will not yield to the data and methods employed here. The empirical regularities so far uncovered merely point the way of maximum interest for future theorizing and research. Ideally, one would wish to ground both the demand and the supply side of the labour market on an explicit theory of optimal choice. The next step, as noted above, would be to specify various hypotheses making the theoretical parameters functions of the observable characteristics displayed by individuals and firms. Such hypotheses would be no less ad hoc than the ones tested here, but they would enter the analysis on a higher theoretical plane and would thus be more readily interpretable using concepts familiar to economists. To test such a model, one would need microdata not only on individuals but also on the firms which employ them. By this means, it should be possible to distinguish supply and demand influences with much greater certainty than has been established here. It is to be hoped that data sets of the kind mentioned become available in Canada before too long.

REFERENCES

Abbott, Michael, and Ashenfelter, Orley. "Labor Supply, Commodity Demand, and the Allocation of Time." Review of Economic Studies, XLIII (October, 1976), 389-412.

Alchian, Armen. "The Rate of Interest, Fisher's Rate of Return over Cost, and Keynes' Internal Rate of Return." American Economic Review, XLV (December, 1955), 938-943.

Arrow, Kenneth J. "Higher Education as a Filter." Journal of Public Economics, II (July, 1973),

Ashenfelter, Orley, and Heckman, James. "The Estimation of Income and Substitution Effects in a Model of Family Labor Supply." Econometrica, XLII (January, 1974), 73-85.

Ashenfelter, Orley, and Johnson, George E. "Unionism, Relative Wages, and Labor Quality in U.S. Manufacturing Industries." International Economic Review, XIII (October, 1972), 488-508.

Atkinson, A.B. "On the Measurement of Inequality." Journal of Economic Theory, VI (September, 1970), 244-263.

Barzel, Yoram. "The Determination of Daily Hours and Wages." Quarterly Journal of Economics, LXXXVII (May, 1973), 220-238.

________, and McDonald, Richard J. "Assets, Subsistence, and the Supply Curve of Labor." American Economic Review, LXIII (September, 1973), 621-633.

Becker, Gary S. "Underinvestment in College Education." American Economic Review, L (MAy, 1960), 346-354.

________. "Investment in Human Capital: A Theoretical Analysis." Journal of Political Economy, LXX (October, 1962, Supplement), 9-49.

________. Human Capital. New York: National Bureau of Economic Research, 1964.

________, and Chiswick, Barry R. "Education and the Distribution of Earnings." American Economic Review, LVI (May, 1966), 358-369.

Becker, Gary S. Human Capital and the Personal Distribution of Income. Ann Arbor: University of Michigan, 1967.

Ben-Porath, Yoram. "The Production of Human Capital and the Life Cycle of Earnings." Journal of Political Economy, LXXV (August, 1967), 352-365.

________. "The Production of Human Capital over Time." In Education, Income, and Human Capital. Edited by W. Lee Hansen. New York: National Bureau of Economic Research, 1970.

Bertram, Gordon. The Contribution of Education to Economic Growth. Economic Council of Canada, Staff Study No. 12. Ottawa: Queen's Printer, 1966.

Blaug, Mark. "The Empirical Status of Human Capital Theory: A Slightly Jaundiced Survey." Journal of Economic Literature, XIV (September, 1976), 827-855.

Blinder, Alan S. and Weiss, Yoram. "Human Capital and Labor Supply: A Synthesis." Journal of Political Economy, LXXXIV (June, 1976), 449-472.

Bluestone, Barry, Murphy, William M., and Stevenson, Mary. Low Wages and the Working Poor (Ann Arbor: Institute of Labor and Industrial Relations, University of Michigan, Mayne State University, 1973).

Bloch, Farrel E., and Smith, Sharon P. "Human Capital and Labor Market Employment." Journal of Human Resources, XII (Fall, 1977), 550-560.

Boskin, Michael J. "The Economics of Labor Supply." In Income Maintenance and Labor Supply. Edited by Glen G. Cain and Harold W. Watts. Chicago: Rand McNally College Publishing Company, 1973.

Bowles, Samuel. "Aggregation of Labor Inputs in the Economics of Growth and Planning: Experiments with a Two-Level C E S Function." Journal of Political Economy, LXXVIII (January/February, 1970), 68-81.

________. "Schooling and Inequality from Generation to Generation." Journal of Political Economy, LXXX (May/June, Supplement, 1972), S219-S251.

Brown, Charles. "A Model of Optimal Human-Capital Accumulation and the Wages of Young High School Graduates." Journal of Political Economy, LXXXIV (April, 1976), 299-316.

Burtless, Gary, and Hansman, Jerry A. "The Effect of Taxation on Labor Supply: Evaluating the Gary Negative Income Tax Experiment." Journal of Political Economy, LXXXVI (December, 1978), 1103-1130.

Cain, Glen G., Freeman, Richard B., and W. Lee Hansen. Labor Market Analysis of Engineers and Technical Workers. Baltimore: Johns Hopkins Press, 1973.

Canada. Department of National Revenue. 1972 Taxation Statistics. Ottawa: Information Canada, 1973.

Canada. Dominion Bureau of Statistics. Principal Taxes and Rates: Federal, Provincial, and Selected Municipal Governments, 1970. Ottawa: Queen's Printer, 1970.

Canada. Economic Council of Canada. Second Annual Review. Ottawa: Queen's Printer, 1965.

Canada. Health and Welfare Canada. Characteristics of Low-Wage Earners in Canada. Social Security Research Report No. 01. Ottawa: Information Services, Department of National Health and Welfare, September, 1976.

Canada. Statistics Canada. Dictionary of the 1971 Census Terms. Ottawa: Statistics Canada, 1972.

________. Economic Returns to Education in Canada. Ottawa: Information Canada, 1974.

________. 1971 Census of Canada. Vol. III, pt. 6. Income of Individuals. Catalogue no. 94-768 (Ottawa: Statistics Canada, May, 1975).

________. Occupational Classification Manual, Census of Canada, 1971. Ottawa: Information Canada.

________. Public Use Sample Tapes: User Documentation. Ottawa: Statistics Canada, 1975.

Chiswick, Barry R. Income Inequality: Regional Analyses within a Human Capital Framework. New York: Columbia University Press, 1964.

________, and Mincer, Jacob. "Time-Series Charges in Personal Income Inequality in the United States from 1939, with Projections to 1985." Journal of Political Economy, LXXX (May/June, 1972, Supplement), 34-66.

Chiswick, Barry, and Mincer, Jacob. "Time-Series Changes in Personal Income Inequality in the United States from 1939 with Projections to 1985 [Reply to Schwartz and Parsons]." Journal of Political Economy, LXXXII (September/October, 1974), 1033-1934.

Comay, Y, Melnick, A, and Pollaschek, M.A. "The Option Value of Education and the Optimal Path of Investment." International Economic Review, XIV (June, 1973), 421-435.

Da Vanzo, Julie, De Tray, Dennis N., and Greenberg, David H. "The Sensitivity of Male Labor Supply Estimates to Choice of Assumptions." Review of Economica and Statistics, LVIII (August, 1976), 313-325.

Denison, Edward. The Sources of Economic Growth in the United States and the Alternatives before Us. New York: Committee for Economic Development, 1962.

Dhrymes, Phoebus J. "Small Sample and Asymptotic Relations between Maximum-Likelihood and Three-Stage Least-Squares Estimators." Econometrica, XLI (March, 1973), 357-364.

Diewert, W. Erwin. "Choice on Labor Markets and the Theory of Allocation of Time." Unpublished discussion paper. Canada. Department of Manpower and Immigration, 1971.

Dodge, David A. Returns to Investment in University Training: The Case of Canadian Accountants, Engineers, and Scientists. Kingston, Ontario: Industrial Relations Centre, Queen's University, 1972.

Donaldson, David, and Eaton, B. Curtis. "Firm-Specific Human Capital: A Shared Investment or Optimal Entrapment." Canadian Journal of Economics, IX (August, 1976), 462-472.

Duncan, Greg J. "Earnings Functions and Nonpecuniary Benefits." Journal of Human Resources, XI (Fall, 1976), 462-483.

Dunlop, John T. "Productivity and Wage Structure." In Income, Employment, and Public Policy. New York: W.W. Norton & Co. Inc., 1948.

Eckaus, R.S. "Estimation of the Returns to Education with Hourly Standardized Earnings." Quarterly Journal of Economics, LXXXVII (February, 1973), 121-131.

Freeman, Richard B. The Market for College-Trained Manpower. Cambridge, Massachusetts: Howard University Press, 1971.

Garbarino, Joseph. "A Theory of Interindustry Wage Structure Variation." Quarterly Journal of Economics, LXIV (May, 1950), 283-305.

Gintis, Herbert. "Education, Technology, and the Characteristics of Worker Productivity." American Economic Review, LXI (May, 1971), 266-279.

Goodman, Leo A. "On the Exact Variance of Products." Journal of the American Statistical Association, LV (December, 1960), 708-713.

Griliches, Zvi. "Notes on the Role of Education in Production Functions and Growth Accounting." In Education, Income, and Human Capital. Edited by W. Lee Hansen. New York: National Bureau of Economic Research, 1970.

________. "Estimating the Returns to Schooling: Some Econometric Problems." Econometrica, XLV (January, 1977), 1-22.

________, and Mason, M. "Education, Income, and Ability." Journal of Political Economy, LXXX (May/June, Supplement, 1972), S74-S103.

Gronau, Reuben. "The Intrafamily Allocation of Time: The Value of the Housewives' Time." American Economic Review, LXII (September, 1973), 634-651.

________. "Wage Comparisons--A Selectivity Bias." Journal of Political Economy, LXXXII (November/ December, 1974), 1119-1143.

Haley, William J. "Human Capital: The Choice between Investment and Income." American Economic Review, LXIII (December, 1973), 929-944.

________. "Estimation of the Earnings Profile from Optimal Human Capital Accumulations." Econometrica, XLIV (November, 1976), 1223-1238.

Hall, Robert E. "Wages, Income, and Hours of Work in the U.S. Labor Force." In Income Maintenance and Labor Supply. Watts, Chicago: Rand McNally College Publishing Company, 1973.

Hanoch, Giora. "An Economic-Analysis of Earnings and Schooling." Journal of Human Resources, II (Fall, 1967), 310-329.

________. Hours and Weeks in the Theory of Labor Supply. Santa Monica: Rand Corporation, August, 1976. R-1787-HEW.

Hanoch, Giora. A Multivariate Model of Labor Supply: Methodology for Estimation. Santa Monica: Rand Corporation, September, 1976. R-1869-HEW.

Hansen, W. Lee. "Total and Private Rates of Return to Investment in Schooling." Journal of Political Economy, LXXI (April, 1963), 128-140.

Hause, John C. "Earnings Profile: Ability and Schooling." Journal of Political Economy, LXXX (May/June, Supplement, 1972), S108-S138.

________. "The Risk Element in Occupational and Educational Choices: Comment." Journal of Political Economy, LXXXII (July/August, 1974), 803-808.

Haworth, C.T., and Rasmussen, D.W. "Human Capital and Inter-Industry Wages in Manufacturing." Review of Economics and Statistics, LIII (November, 1971), 376-380.

Heckman, James. "Life Cycle Consumption and Labor Supply: An Explanation of the Relationship between Income and Consumption over the Life Cycle." American Economic Review, LXIV (March, 1974), 188-194.

________. "A Life-Cycle Model of Earnings, Learning, and Consumption." Journal of Political Economy, LXXXIV (August, 1974), S11-S44.

________. "Estimates of the Human Capital Production Function Embedded in a Life-Cycle Model of Labor Supply." In Household Production and Consumption. Edited by Nestor E. Terleckyj. New York: National Bureau of Economic Research, 1975.

Hood, W., and Rees, R.D. "Inter-Industry Wage Levels in United Kingdom Manufacturing." Manchester School of Economic and Social Studies, XLII (June, 1974), 171-183.

Houthakker, Hendrik S.,"Education and Income." Review of Economics and Statistics, XLI (February, 1959), 14-17.

Ishikawa, Tsuneo. "Family Structures and Family Values in the Theory of Income Distribution." Journal of Political Economy, LXXXIII (October, 1975), 987-1008.

Johnson, George E. "The Demand for Labour by Educational Category." Southern Economic Journal, XXXVII (October, 1970), 190-204.

________, and Youmans, Kenwood C. "Union Relative Wage Effects by Age and Education." Industrial and Labor Relations Review, XXIV (January, 1971), 171-179.

Johnson, Harry G. "Some Micro-Economic Reflections on Income and Wealth Inequalities." Annals of the American Academy of Political and Social Science, CDIX (September, 1973), 54-59.

Johnston, J. Econometric Methods. Second edition. New York: McGraw-Hill Book Company, 1972.

Johnson, Thomas. "Returns from Investment in Human Capital. American Economic Review, LX (September, 1970), 546-560.

________, and Hebein, Frederick J. "Investment in Human Capital and Growth in Personal Income, 1956-1966." American Economic Review, LXIV (September, 1974), 604-615.

Jorgenson, Dale W. "The Theory of Investment Behavior." Determinants of Investment Behavior. Edited by Robert Ferber. New York: Columbia University Press for the National Bureau of Economic Research, 1967.

Juster, F. Thomas. "Introduction and Summary." In Education, Income and Human Behavior. Edited by F. Thomas Juster. New York: McGraw-Hill Book Company, Inc., 1975.

Kerr, Clark. "The Balkanization of Labor Markets." In E. Wight Bakke, et al., Labor Mobility and Economic Opportunity. New York: Technology Press and John Wiley and Sons, Inc., 1954.

Kiker, Bernard F. "The Historical Roots of the Concept of Human Capital." Journal of Political Economy, LXXIV (October, 1966), 481-499.

Klevmarken, Anders, and Quigley, John M. "Age, Experience, Earnings, and Investments in Human Capital." Journal of Political Economy, LXXXIV (February, 1976), 47-72.

Knapp, Charles B., and W. Lee Hansen. "Earnings and Individual Variations in Postschool Human Investment." Journal of Political Economy, LXXXIV (April, 1976), 351-358.

Kosters, Marvin. "Effects of an Income Tax on Labor Supply." In The Taxation of Income from Capital. Edited by Arnold C. Harberger and Martin J. Bailey. Washington: The Brookings Institution, 1969.

Kuch, Peter, and Haessel, Walter. An Analysis of Earnings in Canada, Ottawa: Statistics Canada, 1979. Catalogue no. 99-758E.

Layard, Richard, and Psacharopoulos, George. "The Screening Hypothesis and the Returns to Education." Journal of Political Economy, LXXXII (September/October, 1974), 985-998.

Lazear, Edward. "Age, Experience, and Wage Growth." American Economic Review, LXVI (September, 1976), 548-558.

Leibowitz, Arleen. "Years and Intensity of Schooling Investment." American Economic Review, LXVI (June, 1976), 321-334.

Leuthold, Jane H. "An Empirical Study of Formula Income Transfers and the Work Decisions of the Poor." Journal of Human Resources, III (Summer, 1968), 312-323.

Levi, Maurice D. "Errors in the Variables Bias in the Presence of Correctly Measured Variables." Econometrica, XLI (September, 1973), 985-986.

Levinson, Harold M. "Unionism, Concentration, and Wage Changes: Toward a Unified Theory." Industrial and Labor Relations Review, XX (January, 1967), 198-205.

Lillard, Lee. "Estimation of Permanent and Transitory Responses in Panel Data: A Dynamic Labor Supply Model." Unpublished report. Santa Monica: RAMD Corporation, 1977.

Love, R., and Wolfson, M.C. Income Inequality: Statistical Methodology and Canadian Illustrations. Ottawa: Statistics Canada, 1976.

Lucas, Robert E.B. "Hedonic Wage Equations and Psychic Wages in the Returns to Schooling." American Economic Review, LXVII (September, 1977), 549-558.

Lydall, Harold F. Structure of Earnings. Oxford University Press, 1968.

Madansky, Albert. "On the Efficiency of Three-Stage Least-Squares Estimation." Econometrica, XXXII (January-April, 1964), 51-56.

Malkiel, Burton G., and Malkiel, Judith A. "Male-Female Pay Differentials in Professional Employment." American Economic Review, LXIII (September, 1973), 693-705.

Marin, Alan, and Psacharopoulos, George. "Schooling and Income Distribution." Review of Economics and Statistics, LVIII (August, 1976), 332-338.

Masters, Stanly H. "Wages and Plant Size: An Interindustry Analysis." Review of Economics and Statistics, LI (August, 1969), 341-345.

Mehmet, Ozay. "Economic Returns on Human Capital Formation in Canadian University Education, 1961 to 1972." University of Windsor. Unpublished (May, 1976).

Miller, Herman P. "Annual and Lifetime Incomes in Relation to Education." American Economic Review, L (December, 1960), 962-986.

Mincer, Jacob. "Investment in Human Capital and Personal Income Distribution." Journal of Political Economy, LXVI (August, 1958), 281-302.

________. "On-the-Job-Training: Costs, Returns, and Some Implications." Journal of Political Economy, LXX (October, Supplement, 1962), 50-79.

________. "The Distribution of Labor Incomes: A Survey." Journal of Economic Literature, VII (March, 1970), 1-28.

________. "Comment [on Ben-Por th]." In Education, Income, and Human Capital. Edited by W. Lee Hansen. New York: National Bureau of Economic Research, 1970.

________. "Education, Experience, and the Distribution of Earnings and Employment: An Overview." In Education, Income, and Human Behavior. Edited by F. Thomas Juster. New York: McGraw-Hill Book Company, 1975.

________. Schooling, Experience, and Earnings. New York: National Bureau of Economic Research, 1974.

________, and Polachek, Solomon. "Family Investments in Human Capital: Earnings of Women." Journal of Political Economy, LXXXII (March/April, Part II, 1974), S76-S108.

Moreh, J. "Investment in Human Capital over Time." Manchester School of Economic and Social Studies, XLV (June, 1977), 141-161.

Nagatani, Keizo. "Life-Cycle Saving: Theory and Fact." American Economic Review, LXII (June, 1972), 344-353.

O'Connell, John F. "The Labor Market for Engineers: An Alternative Methodology." Journal of Human Resources, VII (Winter, 1972), 71-86.

Oulton, Nicholas. "The Distribution of Education and the Distribution of Income." Economica, XLI (November, 1974), 387-402.

Paglin, Morton. "The Measurement and Trend of Inequality: A Basic Revision." American Economic Review, LXV (September, 1975), 598-609.

Parsons, Donald O. "Specific Human Capital: An Application to Quit Rates and Layoff Rates." Journal of Political Economy, LXXX (December, 1972), 1120-1143.

________. "The Cost of School Time, Foregone Earnings, and Human Capital Formation." Journal of Political Economy, LXXXII (March/April, 1974), 251-266.

Podoluk, Jenny R. Incomes of Canadians. Ottawa: Dominion Bureau of Statistics, 1968.

________. "Some Comparisons of the Canadian-U.S. Income Distributions." Review of Income and Wealth, XVI (September, 1970), 279-302.

Polachek, Solomon W. "Differences in Expected Post-School Investment as a Determinant of Market Wage Differentials." International Economic Review, XVI (June, 1975), 451-470.

Psacharopoulos, George. "Schooling, Experience and Earnings: The Case of an LDC." Journal of Development Economics, IV (March, 1977), 39-48.

________, and Hinchliffe, Keith. Returns to Education: An International Comparison. Amsterdam: Elsevier Scientific Publishing Company and San Francisco: Jossey-Bass Inc., 1973.

Raynor, A.C. "On the Identification of the Supply Curve of Working Hours." Oxford Economic Papers, XXI (July, 1969), 293-298.

Reder, Melvin W. "Gary Becker's Human Capital: A Review Article." Journal of Human Resources, II (Winter, 1967), 97-104.

Rees, Albert, and Shultz, George P. Workers and Wages in an Urban Labor Market. Chicago: University of Chicago Press, 1970.

Reynolds, Lloyd G. Structure of Labor Markets. New York: Harper and Brothers, Inc., 1951.

Riley, John G. "Information, Screening, and Human Capital." American Economic Review, LXVI (May, 1976), 254-260.

Rosen, Harvey S. "Taxes in a Labor Supply Model with Joint Wages-Hours Determination." Econometrica, XLIV (May, 1976), 485-508.

Rosen, Sherwin. "Learning and Experience in the Labor Market." Journal of Human Resources, VII (Summer, 1972), 326-345.

________. "Measuring the Obsolescence of Knowledge." In Education Income, and Human Behavior. Edited by F. Thomas Juster. New York: McGraw-Hill Book Company, 1975.

________. "Human Capital: A Survey of Empirical Research." University of Rochester, unpublished discussion paper no. 76-2, 1976.

________, and Welsh, Finis. "Labor Supply and Income Redistribution." Review of Economics and Statistics, LIII (August, 1971), 278-282.

Schultz, Theodore W. "Investment in Human Capital." American Economic Review, LI (March, 1961), 1-17.

________. "The Rate of Return in Allocating Investment Resources to Education." Jorunal of Human Resources, II (Fall, 1967), 293-309.

________. "The Reckoning of Education as Human Capital." In Education Income, and Human Capital. Edited by W. Lee Hansen. New York: National Bureau of Economic Research, 1970.

Schwartz, Aba, and Parsons, Donald O. "Time Series Changes in Personal Income Inequality in the United States from 1939 with Projections to 1985: A Comment." Journal of Political Economy, LXXXII (September/October, 1974), 1027-1032.

Segal, Martin. "The Relation between Union Wage Impact and Market Structure." Quarterly Journal of Economics, LXXVIII (February, 1964), 96-114.

Sheshinski, Eytan. "On the Individuals' Lifetime Alocation between Education and Work." Metroeconomica, XX (January, 1968), 42-49.

Shishko, Robert, and Rostker, Bernard. "The Economics of Multiple Job Holding." American Economic Review, LXVI (June, 1976), 298-308.

Slichter, Sumner H. "Notes on the Structure of Wages." Review of Economics and Statistics, XXXII (February, 1950), 80-91.

Smith, Adam. An Inquiry into the Nature and Causes of the Wealth of Nations. Edited by Edwin Cannan. New York: Modern Library, 1937.

Stiglitz, J.E. "The Theory of Screening, Education, and the Distribution of Income." American Economic Review, LXV (June, 1975), 283-300.

Taubman, Paul J. Sources of Inequality in Earnings. Amsterdam: North-Holland Publishing Company, and New York: American Elsevier Publishing Company, Inc., 1975.

________, and Wales, Terence. "The Inadequacy of Cross-Section Age-Earnings Profiles when Ability is Not Held Constant." Annals of Economic and Social Measurement, I (July, 1972), 363-370.

________,and same person as above."Higher Education, Mental Ability, and Screening." Journal of Political Economy, LXXXI (January/February, 1973), 28-55.

Theil, Henri. Economic Forecasts and Policy. Amsterdam. North-Holland Publishing Company, 1961.

Thurow, Lester. "Comment [on Ben-Porath]." In Education, Income, and Human Capital. Edited by W. Lee Hansen. New York: National Bureau of Economic Research, 1970.

Tolley, G.S., and Olsen, E. "The Interdependence between Income and Education." Journal of Political Economy, LXXIX (May/June, 1971), 460-480.

United States. Bureau of the Census. Statistical Abstract of the United States, 1970. Washington: U.S. Government Printing Office, 1970.

Wachtel, Paul. "The Returns to Investment in Higher Education: Another View." In Education, Income, and Human Behavior. Edited by F. Thomas Juster. New York: McGraw-Hill Book Company, 1975.

Wachter, Michael L. "Relative Wage Equations for United States Manufacturing Industries, 1947-1967." Review of Economics and Statistics, LII (November, 1970), 405-410.

Wales, Terence J. "Estimation of a Labour Supply Curve for Self-Employed Business Proprietors." International Economic Review, XIV (February, 1973), 69-80.

________, and Woodland Alan D. "Labour Supply and Progressive Taxes." Review of Economic Studies, XLVI (January, 1979), 83-95.

Wallace, T.D., and Ihnen, L.A. "Full-Time Schooling in Life-Cycle Models of Human Capital Accumulation." Journal of Political Economy, LXXXIII (February, 1975), 137-156.

Weisbrod, Burton A. "The Valuation of Human Capital." Journal of Political Economy, LXIX (October, 1961), 425-436.

Weiss, Leonard W. "Concentration and Labor Earnings." American Economic Review, LVI (March, 1966), 96-117.

Weiss, Yoram. "Investment in Graduate Education." American Economic Review, LXI (December, 1971), 833-852.

Welch, Finis. "Education in Production." Journal of Political Economy, LXXVIII (January/February, 1970), 35-59.

________. "Human Capital Theory: Education, Discrimination, and Life Cycles." American Economic Review, LXV (May, 1975), 63-73.

Wilkinson, Bruce W. "Present Values of Lifetime Earnings for Different Occupations." Journal of Political Economy, LXXIV (December, 1966), 556-572.

Zellner, Arnold, and Theil, Henri. "Three-Stage Least-Squares: Simultaneous Estimation of Simultaneous Equations." Econometrica, XXX (January, 1962), 54-78.