A HUMAN CAPITAL APPROACH TO
OCCUPATIONAL WAGE DIFFERENTIALS

by

ANDREW LESLIE ROBB

B.A., University of British Columbia, 1966

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF ARTS

in the Department
of
Economics

We accept this thesis as conforming to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA
September 1967
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
Vancouver 8, Canada

Date September 18, 1967
The occupational wage structure has been a subject of much interest in recent years. The interest, however, has concentrated primarily on the short-run aspects of the problem much to the neglect of the long-run. The recent interest in investment in education has prompted this theoretical and empirical study of the long-run occupational earnings structure from the point of view of investment in education.

The paper begins by constructing a theoretical model of occupational earnings in which the earnings of an occupation are related to the investment in formal and informal (on the job training and learning by doing) education associated with that occupation. The relation between various occupational earnings streams is established by equating the present values of the expected earnings streams of all occupations. From the theoretical relationship, it can be predicted that the functional relationship between earnings and education should be non-linear with first and second derivatives positive. Moreover, it can be predicted that the degree of non-linearity will be related to the rate of discount that is applied to investment in education. From this section of the paper arise two important conclusions for studies of long-run changes in the occupational wage structure.
Firstly, in studying long-run changes in the occupational earnings structure, attention must be paid to the changing distribution of investment in education among the occupations. Secondly, a change in the shape of the functional relation between earnings and education could be related to long-run changes in the appropriate discount rate.

The empirical section of the paper tests the predicted relation between earnings and education by means of regression analysis. The prediction that the relationship should be non-linear (first and second derivatives positive) is borne out by these tests. Moreover, the degree of non-linearity in the empirical relation appears to be approximately the same as predicted from the theoretical model. Finally, using only simple measures of schooling, over 80% of the occupational earnings structure could be explained in the regression analysis.
TABLE OF CONTENTS

ABSTRACT .................................................. 1
TABLE OF CONTENTS ....................................... iii
LIST OF FIGURES ........................................ iv

CHAPTER

I  INTRODUCTION .......................................... 1
II  THE THEORETICAL MODEL .............................. 5
III  PRACTICAL SIGNIFICANCE AND EMPIRICAL
    SPECIFICATION OF THE MODEL ......................... 26
    III.A. The Practical Significance of the model ....... 27
    III.B. The Empirical Specification .................. 35
IV  EMPIRICAL TESTING .................................... 44
V  CONCLUSION AND SUMMARY ............................ 50

BIBLIOGRAPHY ............................................... 55

APPENDIX I .................................................. 58
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>II-1</td>
<td>Earnings-Schooling Relation</td>
<td>9</td>
</tr>
<tr>
<td>II-2</td>
<td>Alternate Life-Earnings Paths</td>
<td>22</td>
</tr>
<tr>
<td>III-1</td>
<td>Graph of the Earnings-Schooling Relation</td>
<td>29</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

At least since the time of Adam Smith, economists have been interested in the occupational choice of individuals and specifically in the differentials in earnings among occupations. The classical economists had a good deal to say about the determinants of the occupational choice and consequently about the relative earnings in occupations. In a recent paper, Melvin Reder has distinguished between the long run and the short run aspects of occupational wage differentials. It is his contention that "economic theorists have tended to explain occupational differences in wages by differences in costs of training,...for long-run differences." In the short-run, demand is held to play a much more important role

---

1 A very good summary of the classical position on this point has been made by S. Rottenberg in "On Choice in Labour Markets," Industrial and Labour Relations Review, IX, (Jan. 1956), pp.183-199.


3 Ibid.
in the determination of occupational differentials. In keeping with other fields of economics, the classicists were more interested in the long-run determinants, while work in this century (at least theoretical work) has been more concerned with short-run problems.

While economists have held the long-run view outlined above, a systematic long-run theory based on such premises has not been worked out. It will be the aim of this paper to provide a formal long-run theory of occupational wage differentials.

The recent interest in investment in education has prompted this study of the theoretical and empirical aspects of occupational wage differentials. More specifically, Gary Becker's theoretical analysis of investment in human capital in which he makes occasional reference to the implications of his analysis for occupational earnings patterns, has shown need for more systematic and thorough research on this point. In addition to Becker's work, from the empirical side, Weiss' article has emphasized the importance of education and training as determinants of wage differentials. These works have served as a starting point for the present project.

---

1 See: M. Reder, op. cit., for a review of recent work.


3 Ibid., p. 39 ff.

It is a thesis of the present work that the distinguishing characteristic of an occupation is that persons in an occupation differ from those in other occupations by education, training, and other learned characteristics.\(^1\) That differences in occupational earnings would arise does not follow directly from this alone, however. It is the differing amounts of time and effort -- which must be compensated for by a satisfactory return -- in acquiring the necessary skills for the various occupations that cause an occupational earnings structure to emerge. Differences arise "because earnings tend to be net of investment costs and gross of investment returns."\(^2\)

In this paper a model is built which treats the different amounts of schooling in the different occupations as alternative investments which will pay equal rates of return. The equilibrium earnings structure is determined by the equalization of present values of alternative earnings streams. By equating two earnings streams of different amounts of schooling, a ratio can be found which relates the constant annual earnings\(^3\) of two occupations which differ in the

---

\(^1\)In practice, to distinguish between occupations is a task parallel to that of delineating a labour market.

\(^2\)G. S. Becker, op.cit., p.43.

\(^3\)A constant annual earnings stream refers to an income which is the same each year - constant over an individual's life.
amount of schooling required by the members. At a later stage in the analysis, earnings streams which increase over a lifetime (in some systematic fashion) are equated and a similar relation is found between the average annual earnings in the two occupations. By this method, a solution can be found which relates average earnings to the number of years of education and training required.\textsuperscript{1}

This, in brief, is the model which is built in Chapter II of this paper. Chapter III goes on to discuss the shape of the relationship and the effects of changes in the various parameters on that shape. Also, in Chapter III, is a discussion of data that can be used to test the model. Chapter IV describes tests that were undertaken in this connection and the results of those tests. The conclusion, Chapter V, discusses the applicability and importance of the model to studies of the occupational wage structure.

\textsuperscript{1}It should be pointed out that since we are interested in training and education which reduces income in the present and increases it in the future, our interest is with 'general' as opposed to 'specific' training. Becker makes this distinction in his article quoted above (p.12-13) which is briefly: general training is marketable to more than one firm while specific training is restricted to one firm -- consequently, the person receiving general training tends to pay for it himself, while specific training is generally paid for by the firm and recaptured later. There are some exceptions to this which Becker points out but which are not very important for this work.
CHAPTER II

THE THEORETICAL MODEL

In the introduction, the theoretical relationship between occupations and earnings (through differences in investment in education) was briefly outlined. In this chapter the basic model of occupational earnings structure will be constructed. This model draws heavily on the work of J. Mincer\textsuperscript{1} and G. Becker\textsuperscript{2} as a starting point.

At the outset, a number of seemingly restrictive assumptions will be made in order to establish the basis of the model. Later, some of these assumptions can be relaxed and incorporated into the model.

Let us begin with the following assumptions:

(1) All persons spend the same number of years involved in education and working although the distribution of time


between the two varies widely.¹

(2) The only costs of education to the individual are foregone earnings -- tuition fees and incidental costs are assumed non-existent.

(3) Education is undertaken only when a 'normal' return can be earned on the investment² -- there are no consumption aspects to education at the time of the investment or in the future.

(4) There are no systematic occupational preferences.

(5) Risk does not enter the occupational decision.

(6) The labour force is composed of individuals of equal ability -- both as learners and producers.

(7) The only method of improving the productivity of an individual is by investment in education at a formal institution and there is no depreciation on this investment. In

¹ Later, this is taken to be between the ages of 12 and 65. After age 12 all education is taken to be an investment and individuals stay in the labour force until age 65. The term 'labour force life' will be used to refer to this total of working and schooling years.

² In this system, the decision process can be envisioned to operate in either or both of the following ways. Each individual at age 12 (the age at which investment begins) can make his occupational and educational choice. However, a decision (to stop or continue investing in education) made at any time during the education process (which may be made on the basis of more information) theoretically is the same as making the decision at age 12. If a person who has studied for N years decides that it would pay a normal return to study two additional years this means that the present values of the two income streams (with or without the additional schooling) will be equated in year N. Further discounting of the two streams back to year zero cannot change that equality.
the first approximation, on the job training and learning by doing do not add to an individual's productivity.

To the above assumptions it is necessary to add the basic economic behavioral statement -- individuals are maximizers. An individual will make the occupational choice that maximizes the present value of the net expected earnings stream than an occupation will yield. Given assumptions 3 to 6 above, the demand forces are unimportant as determinants of the long-run occupational earnings structure, and consequently, equilibrating forces will tend to bring the occupational earnings structure into a long-run equilibrium consistent with a normal return on the necessary investment in education. What follows in this chapter is an analysis of the long-run equilibrium that follows from the above assumptions.

Given the above framework, it is possible to derive an equilibrium relationship between annual earnings and schooling for an individual. This relationship will express the alternative earnings the individual could receive by investing in various amounts of education. Since (by assumption) this relationship is the same for all persons, it will represent the relationship between schooling and earnings for society

---

1 The assumptions made, imply a perfectly elastic long-run supply of workers to each occupation within the relevant range -- leaving the occupational earnings structure to be established by supply considerations only.
as a whole. By the above assumptions and the definition of an occupation (see Chapter I) this relationship can be used to express the occupational earnings structure -- an occupation can be associated with a specific number of years of schooling and consequently with an earnings level.

Since by assumption, an individual's annual earnings are the same in every year, the present value of lifetime earnings (before education commences) for an individual can be expressed as follows;

1. \[ P_V = A_n \int_0^L e^{-rt} \, dt = A_n \frac{e^{-rL} - e^{-rL}}{r} \]

where,

- \( PV_n \) = the present value of an earnings stream for an individual with \( n \) years of schooling,
- \( A_n \) = the annual earnings for an individual with \( n \) years of schooling,
- \( L \) = the length of 'labour force life' -- defined in footnote 1 on page 6,
- \( n \) = the number of years of schooling,
- \( r \) = the discount rate,
- \( t \) = time in years,

and the discounting process is continuous.

Similarly the present value of lifetime earnings for a person with no investment in education can be written as;

2. \[ PV_0 = A_0 \int_0^L e^{-rt} \, dt = A_0 \frac{1.0 - e^{-rL}}{r} \]

---

1 The form of these first two equations is taken from J. Mincer, *op.cit.*, p.285.
Equating the present values of these two income streams (in accordance with the long-run equilibrium condition) the following relationship is found:

3. \[ f(n) = A_n = A_0 \frac{e^{rL} - 1.0}{e^{r(L-n)} - 1.0} \]

It is then possible to plot \( A_n \) against \( n \) to discover the form of the relationship between schooling and earnings.  

FIGURE II-1  
Earnings - Schooling Relation

\(^1\) Equating these streams of earnings assumes a maximizing sort of behavior on the part of workers. This disallows any income-leisure substitution at higher income levels. Theoretically and practically it can be shown that some substitution occurs, but it is impossible to specify the precise nature of this relationship and consequently has been left out of the formal analysis of this thesis.
The relationship plotted in Figure II-1 is concave from above, i.e., both first and second derivatives are positive (at least over the relevant range -- with \( n \) between 0 and \( L \)). This non-linear relationship is of some interest to the economist studying occupational earnings structures. However, the importance attached to this finding will be related to the degree of non-linearity of the relationship. This will be discussed in Chapter III where simulated earnings-schooling functions are studied.

Having formulated the earnings -- education relationship under the restrictive assumptions spelled out at the beginning of this chapter it is advisable to relax some of these assumptions and study the effects of the others on the basic relationship. After this has been completed it will be possible to formulate hypotheses for testing.

The assumptions will be considered in the order they are listed at the beginning of this chapter.

1. The constant 'labour force life.'

The assumption that all persons are in the labour force (which is taken to mean education plus working life) for a specific number of years can be relaxed without creating any problems for the above analysis. The 'labour force life' that is considered here extends between 50 and 60 years. Any extension of this life by a person in the labour force can be considered to yield income 50 to 60 years in the future. The addition to the present value of earnings 50 years in the future is so small as to be negligible. For example, compound
discounting for 50 years at a rate of 10% reduces a sum to less than one per cent of its face value.

An alternative to the above assumption of equal 'labour force lives' for all persons would be to assume infinite working lives for those in the labour force. This latter assumption has been used by others\(^1\) and has been justified on the same grounds as the assumption made in this paper. Relaxing this assumption would not change the previously derived results perceptibly however, for mathematical flexibility the assumption will be kept as stated.

(2) The lack of tuition fees

The second assumption, that of opportunity costs being the only real costs of investment in human capital can be dealt with in much the same manner as the assumption above. Allowing tuition fees and other incidental costs of schooling to exist for the moment, suppose that the total costs of this type were constant for every year of schooling or in fact were increasing. This would not change the shape of the basic relationship previously outlined although this change would tend to make the relationship more non-linear. On the other hand, if tuition fees decreased with increased years of schooling, that is, each additional year of schooling costs

\(^1\) See, for example, G. Becker, *op.cit.*., p.36.
slightly less in the way of tuition fees, this would appear to have the effect of reducing the degree of non-linearity in the original relationship. It turns then to an empirical question of the nature of tuition and incidental costs of schooling. These costs (to the individual of course, not to society) appear to be low in the early years of schooling, rising in the early years of university and falling off or steady in graduate schools (the fees of graduate school are hard to ascertain, for some they are higher, but for some they are almost non-existent, being covered by large fellowships). To assume a specific shape for the cumulative cost curve of schooling is an extremely difficult task. Since these costs comprise a small part of total cost and are so difficult to estimate, the author felt it best to assume a constant cost from year to year — that constant level being zero.

(3) The lack of consumption benefits to education

From discussions devoted to the subject of education and from personal knowledge, it is obvious that many persons feel that there are consumption aspects involved in the education. These benefits apparently take the form of consumption benefits derived while the investment in education is being undertaken and/or consumption benefits derived later in life. The consumption benefits derived at school would tend to narrow earnings differentials in that persons would be willing to undertake investment in education at a lower apparent rate of
return than would otherwise be the case.\(^1\) However, if these
types of consumption benefits do not increase (as a propor-
tion of schooling costs) with increasing years of schooling
they should not affect the previous results.\(^2\) By reason of
ignorance on the part of social scientists, this assumption
appears to be as valid as any.

About consumption benefits derived later in life —
these being of the nature of increased occupational status,
the "enjoyment of a fuller, more complete life," etc.,\(^3\) — it
is difficult to formulate anything that can add to the theo-
retical model. If in fact education is undertaken with the
idea that there will be a non-monetary return to the invest-
ment as well as a monetary one, by the equilibrium assumption,
lower earnings would be observed than would otherwise be
predicted in the occupations requiring large amounts of educa-
tion. It should be recognized that the fact that there are

\(^1\)This is not to say that investment in education would
yield a lower rate of return than other investments in the
economy. It is likely that students who are investing in
education are making no other investment because of lack of
money. For the most part, students are relatively poor and
may therefore demand a high return to entice them to forego
current consumption. That is, because investment in education
is undertaken by a certain group of people, the return on
that investment may not be equalized with other rates in the
market.

\(^2\)That is, consumption benefits may in some way be the
'same' in each year, so that only the rate of return is
affected, not the structure of earnings.

\(^3\)See B. Weisbrod, "Education and Investment in Human
106-123, for a more complete discussion of monetary and non-
monetary returns to education.
post hoc consumption benefits is not in itself sufficient to cause the reduction in earnings differentials just mentioned. It is necessary also that these benefits be foreseen by those investing in education and that these benefits be included in the calculations of the return on the investment.\(^1\) To the degree that this in fact the case, there will be some effect on the basic relationship. The degree of the effect, however, is a quantitative, empirical question that cannot be predicted a priori.

\((4)\) No systematic job preferences

The assumption made about the lack of systematic job preference is intimately linked with the previous assumption about the consumption aspects of education. The previous assumption refers to benefits directly linked with the educational process whereas the present assumption refers to benefits that are job specific (it is a matter of indifference whether these benefits are termed consumption benefits or non-monetary income). As with consumption benefits of education, systematic job preferences would tend to distort the previously discussed relation between education and earnings. The distortion would not tend to be systematic with relation to education, however, since non-monetary

\(^1\)This also is true for consumption benefits of the first nature. It is not enough that social scientists look back on their school days with sentimentality, students must fell that they are gaining consumption benefits while they are investing.
benefits could exist in occupations at any level of education. It is true without question that some individuals derive large non-monetary benefits from specific occupations. However, this fact alone is not enough to require a change in the theory. It is equally true that all persons do not prefer the same occupation or occupations (even though a few occupations are generally liked and a few generally disliked). The basic economic force that enabled a prediction in the first instance was the behavioral assumption of maximizing behavior on the part of individuals. Under this assumption, it is the marginal adjustment process that established the earnings structure. That some individuals in every occupation earn 'rents' cannot alter this marginal process. Thus, while it is clear that individuals do have job preferences, it is not clear (and cannot be predicted a priori) that this phenomenon is significant enough to distort the previously deduced relationship between earnings and education in anything but a random manner.

(5) The lack of risk in the occupational choice

There are at least three reasons why the assumption has been made that risk does not affect the occupational

---

1It is generally thought that such benefits accrue in occupations requiring higher levels of education -- in which case it is difficult to distinguish this assumption from the previous one -- but this is not necessarily the case.
decision. In the first place, since it is occupations that are being considered, it may be claimed that there is little reason to expect individuals to attach more or less risk to different occupations -- that is to say, that individuals would expect an equal chance of success in various occupations (at least with respect to the occupations which lie in his range of abilities). In the second place, it is possible that individuals do not even consider risk when making the occupational decision. This second point may overlap the first in that the reason that risk is not even considered is because individuals 'feel' that the various occupations are 'equally risky'. There may, however, be other reasons that risk is not considered in the occupational decision. In the third place, it may be that risk is considered but that its effects are of various types that tend to cancel out in aggregate. For these reasons, risk has been assumed neutral in this model.  

---

1One of these possible reasons may be that while an individual rationally feels that he should allow for risk in this decision, he has no idea how to estimate risk to bring it into the decision.

2M. Friedman and S. Kuznets in Income from Independent Professional Practice, (New York: National Bureau of Economic Research, 1954), Chapter 3, hypothesize that individuals will act as risk takers because they generally overestimate their own ability. The author has held the opposite view with some success -- that persons will act as risk averters, see "Occupational Wage Differentials and Investment in Human Capital", unpublished paper, 1967. (This is the more usual assumption -- see for examples J. Tobin, "Liquidity Preference as Behavior towards Risk", Review of Economic Studies, XXV (Feb. 1958) pp.65-86).
(6) The ability of individuals

If some specific distribution of ability is allowed for in the model (rather than assuming individuals equal with respect to ability as in the original assumption) there are a number of ramifications for the structure of earnings. The degree of adjustment to the previous model depends on the precise distribution of ability among the various occupations. It should be first recognized that ability cannot be interpreted to be a uni-dimensional concept. One individual may have an ability to be a carpenter and another the ability to be a medical doctor, but to compare these abilities on a uni-dimensionally scale is not logically possible. Each of these individuals may be said to be the best in his field but further comparison (with respect to ability) is absurd. Secondly, it should be pointed out that ability is not generally a narrow or specific characteristic. A person with high ability in one field generally has a similar ability in closely related fields. Thus, the highly able medical doctor referred to above, may have chosen to become a chemist or a dentist (and also excelled in those fields) had the return to the investment in education in those fields been higher (than the return to medical doctors) during the period in which he made his occupational choice.

As a first step, suppose there existed a well-defined set of occupations, each of which contained a sizeable number of workers. Suppose moreover, that each individual has different abilities and the ability of an individual can be
utilized in more than one occupation. In fact, abilities in this labour force are distributed such that no occupation is lacking workers of the required ability to such a degree that it is necessary to pay a rent in that occupation — that is, each occupation pays a normal return to the investment in human capital necessary to allow an individual to enter that occupation. Moreover, it can be added that the distribution of ability to perform a specific occupation is more or less normally distributed (within a specific occupation a uni-dimensional concept of ability is valid, that is, the ability to perform a specific task) within each occupation.

Such a model of the distribution of ability appears to the author as a not unreasonable position to hold.¹ This model does not require that each occupation pay a standard wage — there will be some distribution (approximately normal) within an occupation although obviously not as great a difference as between occupations. However, since the distribution of talent available to an expanding occupation is the same as the distribution of already employed talent in that occupation, the long-run supply curve to an occupation can be considered as horizontal and the return on investment

¹ This position could alternatively be expressed in Longfield's ideas of competition operating through a chain of intermediate professions. For an outline of this view, see E. Whittaker, A History of Economic Ideas (Toronto: Longmans, Green and Co., 1943) p.501.
in education for each occupation will be kept at a 'normal' rate of return. No distortion in the basic model arises if it is altered in this fashion.

While the above reformulation is not an entirely realistic view of the labour market, it is much closer to reality than the initial restrictive assumption. It is obviously true at the upper end of the occupational scale (with respect to the amount of investment required) that some individuals are paid rents for a unique ability but there is at least one reason to discount this as a problem when the above (reformulated) model is applied to the real world. The frequency of occurrence is probably rare — note that a distribution of abilities is already allowed for and it is the deviations from this distribution and not deviations from the 'norm' that will cause distortions in the model.

(7) On the Job Training and Learning by Doing.

The final assumption that must be considered theoretically is the assumption that on the job training and learning by doing do not add to the productivity of an individual. This assumption must be relaxed unless the previous results are to be rendered meaningless. It is well known that on the job training and learning by doing are very important aspects of an individual's investment in education. Until now, this paper has considered only earnings' paths which are constant over an individual's life because of the assumed non-
existence of informal training. Allowing for informal training introduces a number of complexities into the analysis not the least of which is the problem of specifying in what manner informal training should be introduced into the model.

Since there is no good measure of informal education available, it seemed of little value to consider informal and formal education in a general way as inputs in a process and later devise tests on this basis. Moreover, from knowledge of the educational process, the author felt this method an unrealistic approach to the problem. Consequently, the method chosen here has been to postulate a specific relation between formal and informal education and incorporate this into the model previously discussed.

In considering the specification of the relationship between formal and informal education, two possible alternative assumptions suggested themselves; a) formal and informal education are substitutable inputs in the production of a worker, b) because of the nature of the two inputs -- formal and informal education -- the inputs tend to be complementary.

The author's view of the education process has caused him to opt for the second of these possible specifications —

---

1 B. W. Wilkinson, in Studies in the Economics of Education, Ottawa: Economics and Research Branch, Department of Labour, 1965, Chapter 2, converted an American measure of informal education requirements (by occupation) into Canadian measures but did not find this data reliable.
that is, that formal and informal education are complementary inputs and the more formal education that is attained, the more informal education that will follow. The author views formal education as a process whereby the student absorbs the capacity to learn. This education is not directed towards fitting the student to perform a specific task, but to fit the student to learn more complex tasks in the future through informal training. Thus, the more formal education that is attained by a student, the more profitable it will be to pursue extended informal training on the job. A student who makes an occupational choice, at the same time makes a choice as to the amount of formal education that will be necessary to suit him for learning, in the informal manner, the specific tasks associated with the occupation.¹

The assumption made above is the major change in the theoretical framework in this section. However, in order to incorporate this assumption into the model, an additional minor assumption (of a more mechanical nature) must also be made. The pattern of informal training, with respect to timing, must also be specified. It is therefore assumed, that on the job training and learning by doing increase the individual's productivity (and consequently his earnings) in a linear fashion. That is, an individual's productivity

¹Knowledge gained from observation of age-earnings profiles also tends to support the case for complementarity. See G. Becker, op.cit., p.43 and J. Mincer, op.cit., p.292.
increases by the same amount each year until some cut-off point is reached after which his earnings are invariant.  

Thus an individual's life earnings path has the shape shown below by \( G_n \) as opposed to the path \( A_n \) which depicts the life earnings path of an individual with \( n \) years of schooling and no informal training.  

**FIGURE II-2 Alternate Life - Earnings Paths**

---

1 The choice of a realistic cut-off point was not simple because of the lack of adequate data from which to make a decision. The following alternatives were studied (and are discussed later) although it turned out that the choice made little difference to the final results: i) that the increases continue over a certain fraction of the working life, ii) that the increases continue a certain number of years after formal education ceases, iii) that the increases continue until a certain number of years before retirement.

2 These three paths are not meant to represent alternatives. \( A_0 \) is a path with no investment in education. \( A_n \) is a path with \( n \) years investment but not allowing for on the job training. \( G_n \) is a variation on \( A_n \), allowing for on the job training.
The point 'D' in the above diagram represents the age at which informal training (and consequently increases in earnings) ceases. The equation of the line $G_n$ can then be written:

4. $G_n = B_n + S_n (t-n)$ for $n \leq t \leq D$

and

5. $G_n = B_n + S_n (D-n)$ for $D \leq t \leq L$

where;

- $S_n$ represents the slope of the line (and $S$ is an increasing function of $n$ by assumption (B) above)

and

- $B_n$ represents the starting salary after $n$ years of schooling -- $B$ is some fraction of $A$.

The present value of the earnings path $G_n$ will of course be equal to the present value of the earnings path $A_n$ in equilibrium and both will be equal to the present value of the earnings path $A_0$ which represents the earnings path of an individual with no formal or informal training.

Given this starting point, it is desirable to find the average lifetime earnings of an individual with $n$ years of schooling. The model can be summarized in the following equations:

6. $A_0 \int_0^L e^{-rt} \, dt = A_n \int_n^L e^{-rt} \, dt$

which follows from equations 1 and 2 above -- equating the present values of two earnings streams for persons with

---

1 This is desirable because it is one measure that can be estimated from the available data.
different amounts of education (and no informal training).

7. \[ A_n \int_0^L e^{-rt} \, dt = B_n \int_0^L e^{-rt} \, dt + S_n \int_0^D (t-n) e^{-rt} \, dt \\
+ S_n (D-n) \int_D^L e^{-rt} \, dt \]

which equates an earnings stream with \( n \) years of schooling and no on the job training to a stream with \( n \) years of schooling plus on the job training in accordance with the two assumptions just made. That is, equating the streams \( A_n \) and \( G_n \) in figure II-2.

8. \[ \bar{A}_n = B_n + \left( \frac{(S_n/2)(D-n)^2 + S_n(L-D)(D-n)}{L-n} \right) \]

where \( \bar{A}_n \) is the average of the earnings stream \( G_n \).

The above equations (6, 7 and 8) can be solved to yield:

9. \[ \bar{A}_n = A_0 \frac{(B_1 + B_2)}{B_2} - S_n \frac{(B_3 - nB_4 + DB_5 - nB_5)}{B_2} \\
+ (\frac{(D-n)^2}{2} + (L-D)(D-n)) * S_n / (L-n) \]

where:

\[ B_1 = \int_0^n e^{-rt} \, dt \]
\[ B_2 = \int_n^L e^{-rt} \, dt \]
\[ B_3 = \int_n^D te^{-rt} \, dt \]
\[ B_4 = \int_n^D e^{-rt} \, dt \]
\[ B_5 = \int_D^L e^{-rt} \, dt = B_2 - B_4 \]

The solution of the system (given in equation 9 above) specifies a relationship between average earnings and years of schooling when on the job training and learning by doing
are allowed to occur in the above-specified manner. It turns out that this relationship -- complicated as it appears to be -- has the same general shape as the more simple relationship given by equation 3 (that is, without the on the job training). ¹

This chapter has attempted to build a theoretical model of the occupational wage structure. A model has now been constructed but is as yet in a very general form. It will be the task of the next chapter to examine the model more closely by specifying the parameters of the model, and to turn the theoretical model into a form which can be subsequently tested.

¹The first and second derivatives of the expression $A_n$ with respect to $n$ are both positive. The equation would not submit easily to mathematical differentiation so realistic values of the parameters were used and the equation was put into a computer to find this result.
CHAPTER III

PRACTICAL SIGNIFICANCE AND EMPIRICAL SPECIFICATION
OF THE MODEL

The aim of Chapter II was to build a theoretical model of the occupational wage structure that could be subjected to statistical tests. It is nothing new to suggest that earnings of an occupation are somehow related to the skill and training required for the occupation of an individual. References to this idea date back at least to Adam Smith. However, the relationship has been of an intuitive nature and has not been made formally theoretical. A formal model has been constructed in the previous chapter, but before any testing can be undertaken there are two aspects of the model that must be discussed;

(A) At various stages in the last chapter, conclusions were made about the shape of the relationship between earnings and investment in education and about the importance of that shape in empirical analysis. It is necessary therefore to begin this chapter by a more precise examination of the suggested

---

relationship. To do this, it was necessary to employ the technique of computer simulation due to the complexity of the relation (see equation 9 of the last chapter). In so doing, another problem arises. To use the computer simulation method, values must be assigned to the various parameters in the model. If variation of the parameters causes marked variation in the original relationship, then great care must be employed in choosing the parameter values. Thus it will be necessary in this section to also discuss the choice (and the importance of the choice) of the parameters of the model.

(B) Also, before any testing can be undertaken, it will be necessary to relate the theoretical notions already discussed to the available data. From the published data\(^1\), variables must be drawn that correspond to the theoretical variables in the model. This will involve comparison of the theoretical concept of occupations and schooling and their practical counterparts. Corrections will have to be employed to bring the two into correspondence.

Part A — The Practical Significance of the Model.

In order to discuss the simulation of the model at all, it is necessary to inform the reader of the range of values that the main independent variable (years of schooling or

\(^1\)In this case the 1961 Census of Canada served as the source of data.
investment in education) was allowed to take on. The author felt it necessary to the theoretical framework to not begin counting schooling as an investment in education until the age (or level) at which a child could make the decision to remain in school or to leave. It was felt that age 12 and grade 6 were most appropriate for this purpose. Having chosen this point as a base, the number of years of schooling in an occupation will fall between zero and thirteen years (actual schooling minus 6 years). The maximum schooling allowed for (thirteen years) represents twelve years in grade school and seven in University -- this easily accounts for the educational attainment in any of the occupations studied.

Having discussed the choice of the independent variable, it is useful to graph immediately some earnings-education relations for particular sets of parameter values. This will be useful to familiarize the reader with the form of the relation which will be discussed in the next few pages and also as a reference point that can be used in discussions of the parameters.

---

1 At the present the compulsory attendance age varies between 12 and 14 (across Canada) or between grade 6 and 8. As this cut-off point was lower in the past revision of this choice may be necessary;
With the aid of the above diagram it is now possible to discuss the choice of the parameters in the model as to their significance. For each of the parameters, a number of alternative values have been studied to discover if an incorrect

1The following parameter values are used in the above diagram; the discount rate at 12 and 14% for the lower and upper curves respectively; retirement age = 65 (i.e. L, the labour force life = 53); A₀, the base wage = $2000; the age at which salary increments cease = L-15 (15 years before retirement); and the slope -- S -- is chosen at = 30n. The X axis measures years of investment in education, and the Y axis measures the unweighted average earnings of an occupation (in dollars). These are the same paths that are tabulated in Appendix I, Tables I and II.
choice would greatly bias the results that follow. Of the parameters, it appeared that only the choice of the discount rate was important. Some of the other parameters were of slight importance and others of no apparent importance. The following paragraphs discuss these results.

The choice of an appropriate discount rate has been based on private rates of return to investment in education calculated by Becker and others. Although no consensus of opinion on the appropriate rate of return prevails the author has chosen rates of 10, 12, and 14 per cent to study. The discount rate has a very important influence on the shape of the plotted function. A small change in the discount rate (a variation of 2% or so) has a marked positive effect

1Important from the point of view that slight variations in the parameter value would result in large changes in the curvature of the basic education-earnings relation. Generally it was clear what range the parameter values should take but not the exact value.


4W. G. Bowen, op. cit., p. 27 ff.

5Since statistical work has been completed, it has been pointed out that a study on Canadian data suggests the rate to be 15 to 20% -- see footnote in G. Bertram, The Contribution of Education to Economic Growth, (Ottawa: Economic Council of Canada, 1966).
on the second derivative -- that is, an increase in the interest rate increases the degree of non-linearity of the function \(^1\) (Tables I and II of Appendix I show the effect of varying the discount rate only). Two rather important implications follow from this finding; i) if the theoretical relation is to be used to approximate the empirical relation, a very careful choice of the discount rate must be made, ii) if, as some contend, there has been a secular decline in interest rates, and in particular in the rate of return on investment in education, definite changes in the occupational earnings structure are implied. \(^2\)

For a retirement age, ages of 65, 70 and 75 were studies. \(^3\)

\(^1\)Some change in the first derivative is also noted but is of little significance as variation of the slope parameter will bring this back into line.

\(^2\)That is, if the decline in the discount rate represents a series of long-run equilibrium rates of return. If the decline represents only a movement towards an equilibrium, the implication is not at all clear from this model. Historical comparisons of the wage structure which discuss the compression of the wage structure over time would do well to consider the long-run decline in interest rates as an explanatory factor in this connection.

\(^3\)This allows comparison of the effect of postponed retirement of all members of the labour force. One possibility not allowed for by this variation is that different occupations (or persons with different levels of formal education) retire at different ages. It was shown in Chapter II however, that due to the discounting of earnings so far in the future, the effects of such behavior should be minimal. Moreover, to have any effect on the relationship whatever, it would be necessary that a decision on retirement age be made at the time of educational and occupational choice. To hypothesize that individuals would expect any retirement age other than the average for the labour force as a whole would be risky.
Recalling that L represents the 'labour force life' -- the number of years of investment in education plus the number of years working -- then for these three retirement ages, L will take on the values of 53, 58, and 63. Variation in the retirement age results in only slight effects to the earnings-education relation. A five year increase in the retirement age results in increases in both the first and second derivatives but the changes are insignificant (the change in the first derivative is the larger but of little interest here). The reader may judge for himself by consulting Tables I and III of Appendix I which reproduce two education-earnings relationships -- the only difference in which is the variation in the retirement age.

In studying the base level of earnings $A_0$ (the earnings level for an individual with no educational investment -- formal or informal) a range of $1600$ to $2000$ was considered. Variations in this parameter tended to shift the earnings-education relation, of course, but also very minor variations

---

1. This results from setting age 12 as the starting point for investment in education.

2. This large a range was not necessary as it was clear that the obvious choice of this parameter would lie close to $2000$, however, a wider choice was considered in case studies were to be undertaken of earlier or later data.

3. Minor, as used here, refers to a change of the order of that between the earnings' paths in Tables I and III of Appendix I.
in the first and second derivatives were noticed.  

The variation of the cut-off point and the slope should be considered simultaneously because they are chosen in that manner -- once a cut-off point is selected, a slope can be chosen which leads to a realistic range of alternatives. Variation in the cut-off point, with appropriate variation in the slope to maintain realistic values of earnings, resulted in only small changes in the shape of the earnings-education relation and further consideration of this choice is unnecessary at this time.

In the past few pages, a discussion of the importance of various changes in the parameters in the model has been presented; all that remains to be accomplished in the present section of this paper is a discussion of the degree of non-linearity of the education-earnings relationship. In order to test the relevance of the non-linear shape of the curve to the 'real world' a simple test has been employed.

---

That this should occur is not at all clear intuitively, it arises because in equation 9 there are three terms added together on the right hand side of the equation. The first term is by far the largest term and has first and second derivatives positive. The other two terms taken together (added) have first derivative positive and second derivative negative. A change in $A_0$ affects only the first term -- by a fractional increase in earnings at each level of education. Adding a term with a negative second derivative tends to reduce the second derivative in the first term. However, after $A_0$ is increased, the added term (combined second and third terms) has less effect on the total relation.

Cut-off point such as; $n + 25$, $L - 15$, and $(L - n)/2$, were typical of those studies. Slopes to correspond to these cut-off points were of the type; $30n$, $35n$, and $40n$.

3 i.e., of the order noted when the retirement age is varied (see Tables I and III in Appendix I).
Linear regressions have been performed using as the dependent variable the average earnings figures calculated from the theoretical relationship (equation 9) and as the independent variable, the appropriate number of years of schooling (investment in education) in a simple and in a squared form. This procedure was carried out for a number of alternative combinations of values of the parameters. The results were nearly identical regardless of the parameter values. The following two equations are typical of the results of the regressions.

1. \[ \bar{A}_n = 900.0 + 1137.3 n \quad R^2 = .976 \]  
   \[ (51.7) \]

2. \[ \bar{A}_n = 2182.8 + 495.9 n + 49.3 n^2 \quad R^2 = .999 \]

These results indicate that the percentage of the dependent variable explained can not be increased greatly by the addition of the non-linear form of the schooling term. The relationship which has been derived theoretically -- the non-linear form, that is -- does not seem particularly important in practical work. While the non-linear result is

---

1 It should be pointed out, however, that the discount rate, as previously shown, does exhibit a positive relation with the second derivative of the education-earnings relation. Hence, the following two regression equations which employ the largest discount rate considered, 14%, will show the maximum significance of the \( S^2 \) term. Even when the \( n^2 \) term is maximized, the importance of it is very minor.

2 The values of the parameters for this particular set of regressions are as follows; \( r \) (discount rate) = 14%, \( A_0 \) (starting wage for those with no educational investment) = $2000, \( D \) (the cut-off point or age after which no person invested further in informal training = retirement age less 15 years, \( L \) (the labour force life -- previously defined) = 53 years, and the slope of the age-earnings profile was chosen as previously outlined.
interesting theoretically, it may not be important in actual tests made on incomes of occupations, which are influenced by many omitted factors.

Some of the more important of the omitted factors were discussed in Chapter II and, from the arguments put forth there, were felt to be of minor importance. From the arguments of Chapter II and this part of Chapter III, therefore, the prediction emerges that a substantial part of the differences in occupational earnings should be explained by the variables $n$ and $n^2$ (representing schooling) although it remains to be seen whether or not the $n^2$ term will be important in such work.

Part B -- The Empirical Specification

In Chapter II and the first part of this Chapter, a theoretical relationship between schooling in occupations and average earnings in the occupations has been constructed and discussed. Without this relationship it would have been possible to have simply hypothesized a relationship and tested it by means of regression analysis. However, having developed the theoretical relationship, it is now clearer how the variables should be entered in a regression analysis.

In the regression analysis it is obvious that some form of average earnings should be used as the dependent variable and some measure of education as the independent variable. However, there are some problems associated with the empirical testing that must be given consideration at this time. The following are the problems which will be given consideration
at this time; 1) the discrepancy between the theoretical notion of an occupation and the notion used in the available data, 2) the discrepancy between the theoretical measure of earnings (unweighted average earnings in an occupation) and the measure available (weighted average earnings), 3) the problems associated with finding an adequate variable to measure schooling from the data available.

(1) In the theoretical treatment we have treated all occupations similarly in the sense that an individual would enter a chosen occupation and remain there for life. In the 'real world', however, there are many occupational transfers that occur over a person's life. The model would depend on occupational transfers in a disequilibrium (short-run) situation in which transfers would be made to effect equilibrium. It is occupational transfers of another type that create a problem in this model. In many cases in the 'real world' an occupation, so defined, is not an occupation in the sense that it is a lifetime vocation, but is only a 'stepping stone' to another occupation. An obvious example of this is the occupation defined (according to the Occupational Classification Manual, Census of Canada, 1961) as 'Nurses-in-training'. Obviously this is not an occupation in the theoretical sense in which the term was used in Chapter II. It is an educational-training stage to the occupation of 'Nursing'. There are many examples of occupations like this -- although the training stage is not so obvious. Managers of various types, for example, often go through a training program during which
time they are classified as salesmen, accountants, clerks, or any of a number of other occupations. Since occupations are not defined in the 'real world' in the way it was necessary to define them in the theoretical model, it will be necessary in the empirical analysis to take account of this factor. Persons who followed such a circuitous route as the above one would tend to have more informal education than would be predicted by the model and the incomes in 'final' occupations of this type would tend to be higher than otherwise expected. Although it is very difficult to measure this factor, a proxy variable which would tend to measure such influences suggests itself. It was felt by the author that the age at which persons typically entered an occupation would reveal any tendency for one occupation to rely on another for training. Unfortunately, the only source from which such a measure could be drawn was the existing age distributions of occupations. It was felt that a measure that would be relatively free of demand influences and would suit this purpose, was the age below which 10% of the persons employed in an occupation fall — this will be referred to as the 'lower decile age'. Not all demand influences can be avoided in any such measure, but this choice is aimed at minimizing the

1 In order to correspond to the theoretical model, one would want to classify an individual as a manager as soon as his formal education was finished — this of course is impossible in reality.
influence. 1

(2) The average earnings concept utilized in the theoretical model is what is generally called an unweighted average earnings figure — that is, the average is not weighted by the numbers at the various age groups in the occupation. Published data always give (when averages are calculated and published) the average earnings of persons actually engaged in a specific occupation — which is a weighted average earnings figure. If data were available that gave average earnings by age and occupation an unweighted average of the type discussed above could be computed. Unfortunately, however, such data are not available. Consequently, the weighted average earnings of the occupations have been used in the following analysis.

It was possible, fortunately, to conduct a rough test of the validity of using the weighted average as opposed to the unweighted average. From the population sample of the

1 Demand would influence this measure in the following manner. Suppose that an occupation was dying out — for example, due to improved technology, very few persons would enter this occupation and the members of this occupation would tend to be older than in most occupations — this follows from the theory of investment in retraining — the older persons in the occupation would be the least likely to move out as they would not expect to be able to recapture the investment in mobility (and possibly training) that would be required. A measure at the lower end of the age distribution, however, minimizes this difficulty.
1961 Census of Canada data have been published giving average earnings by age groups for selected occupations. In coordination with previous work done by the author a test was conducted on a sample of 20 occupations to find if any significant difference existed between the occupational weighted average earnings and an unweighted average computed from the above mentioned data. A correlation between these two sets of earnings figures yielded a correlation coefficient of .998, revealing that the weighted average approximated the unweighted average extremely well. Moreover, if the unweighted average is regressed on the weighted average as follows:

\[ Yw = K + B*Yu \]

where \( Yw \) = weighted average

---


3These 20 occupations are chosen from the same occupational classification as used later, but since data was given for only selected occupations the choice could not be random.

4The unweighted average used does not correspond exactly with the theoretical model because rather than averaging the (average) earnings at each age, only the average of the (average) earnings for the age groups 25-34, 35-44, 45-54, and 55-64 could be found. For the above mentioned purposes it was felt that this was a reasonable approximation.
Yu = unweighted average

K = regression constant

the regression coefficient B = 1.04. The weighted average earnings figures are just slightly lower than the unweighted figures that would be the more desirable measure. The differences between these two variables are so small as to warrant the use of the weighted average.

(3) The third problem that is encountered when an attempt is made to devise a test of the previously outlined theory is the lack of good data on investment in education. The census provides information on school attendance -- publishing data on the highest grade attended according to the following breakdown; Elementary less than grade 5, Elementary grade 5 and over, Secondary 1 - 2 years, Secondary 3 years, Secondary 4 - 5 years, Some University, University degree. For the purposes of a test on investment in education this data is insufficient in the following respects; (i) the data on schooling in the census do not include schooling undertaken at vocational schools, institutes of technology, and other such institutions. These institutions certainly account for a great deal of investment in education that is missed in the available data.  

One important instance is that in some provinces (and abroad), engineers are trained in technological institutes and not in universities.
what had been labelled the 'lower decile age' variable will tend to catch this investment in education in the regression analysis. To some extent, therefore, this error in the data is not as damaging as it first sounds. (ii) the task of transforming the above-mentioned school achievement groupings into a random variable is somewhat difficult. To assign numerical values to the various groups is a more or less arbitrary decision that is based on the author's (limited) knowledge of schooling in Canada. The following weights have been chosen; Elementary, less than grade 5 = 4, Elementary grade 5 and over = 6, Secondary 1 - 2 years = 8, Secondary 3 years = 10, Secondary 4 - 5 years = 12, Some University = 14, and University degree = 16. (iii) the third problem with the available data is closely related to the second. The values used above, although admittedly arbitrary are acceptable as far as they go. However, in the groupings available from the census, there is no differentiation between various university degrees. While the weight of 16 is appropriate for a bachelor degree it is not at all appropriate for higher degrees. It is not acceptable, moreover, simply to upgrade the weighting used in the last category for all occupations. It is obvious that some occupations call for (almost as a matter of necessity) more university education than a bachelor's degree. For this reason the author has made an attempt to vary the weighting attached to the highest educational grouping according to
the occupations involved. For example, it is well known that an M. D. degree comprises more than 16 years of schooling -- the education required for this degree varies (across Canada) between 18 and 19 years of schooling.\(^1\) In the census data, persons are classified as doctors with less than this amount of education and for this reason the author decided that the best weighting for this group would be 18 years of schooling. Similar corrections (not always the same weighting) have been carried out for other occupations.\(^3\) By this technique, it is felt that a more accurate measure of schooling has been achieved.

From the above discussion, it should be apparent that the theory can be tested in a rather simple form. In the next section regression and correlation analysis will be used to test the data against the theoretical model. Regressions have been undertaken with weighted average earnings as the dependent variable and measures of investment in education as the independent variables. Variables have been incorporated using years of schooling, years of schooling squared, and also previously discussed proxy variable

\(^1\)The same is true of dentists.

\(^2\)Some even who have not completed high school.

\(^3\)The following occupations have been given weightings higher than 16 for the highest educational grouping; Physicians and Surgeons, Dentists, Pharmacists, Veterinarians, Engineers, Scientists, Lawyers, and a few other similar professions.
which has been termed the 'lower decile age' variable. A separate regression was carried out using a theoretically derived variable -- the average earnings that would be predicted by the theoretical model, given the number of years of schooling associated with each occupation. The precise form of the variables, the manner in which they were entered into the regressions, and the results of the analysis will be presented in the next chapter.

---

1The same parameters were used in the derivation of the theoretical variable as were used in the diagram sketched in the Part A of this Chapter -- Figure III-1.
CHAPTER IV

EMPIRICAL TESTING

In Chapter II of this paper, a theoretical framework in which to study occupational wage differentials was outlined. In Chapter III, special problems associated with choosing values for the parameters of the model and with selecting data to conform to the theoretical variables in the model were discussed. Having completed these sections of the work, it is now possible to turn to the empirical testing of the model. The estimated regression equations will first be written down and then discussed in turn.

1) \[ \bar{A}_n = 988.0 + 582.0n \] 
   \[ R^2 = .719 \] 

2) \[ \bar{A}_n = 962.6 + 293.8n + 28.6n^2 \] 
   \[ R^2 = .734 \] 

3) \[ \bar{A}_n = 811.5 + 515.6n + 164.9p \] 
   \[ R^2 = .811 \]
4) \[ \overline{A}_n = -681.0 + 276.2n + 23.9n^2 + 161.1P \]
\[ (64.6) \quad (6.2) \quad (14.5) \]
\[ R^2 = .822 \]

The following are the variables used in the above equations:

- \( \overline{A}_n \) is the weighted average earnings of an occupation, calculated from the 1961 Census.
- \( n \) is a variable representing the number of years of investment in education also calculated from the 1961 Census.
- \( P \) is the proxy variable representing the 'lower decile age' also from the 1961 Census.

These four equations are regressions of the observed variables representing investment in education on the weighted average earnings of the occupations calculated from the census. These regressions reveal that a large proportion of occupational earnings can be explained by these simple measures of investment in education.

\[ R^2 \] is the coefficient of determination and the bracketed numbers below the regression coefficients refer to the standard error of the estimator.

256 observations were employed in these equations.


5Ibid.
An interesting comparison can be made between these regressions and the regressions on the theoretical variables presented at the end of Part A in Chapter III. While the present empirical variables explain less than did the corresponding theoretical variables, the additional amount explained by the inclusion of the schooling in its squared form is very nearly the same in the two cases — the additional amount explained is in the order of 2% and the size of the coefficient of the squared term is about 10% of the coefficient of the linear term in both instances. The degree of non-linearity in the 'real world' relationship of education and earnings appears to bear out the theoretical predictions very well.

In all cases, the regression coefficients show to be statistically significant (with an F-test at the 5% level). The fourth equation is the best estimator of average earnings with over 80% of the variation in occupational earnings being explained — note that the addition of the age of entry proxy variable adds considerably (9 percentage points) to the amount explained of the dependent variable. This

The original prediction was that the long-run occupational wage structure in equilibrium would be explained by these variables. However, the cross section data used here will measure the state of the wage structure at a point in time only — i.e., not likely an equilibrium situation. A reasonably good R² would be predicted nevertheless since the equilibrating forces would keep the wage structure in the 'neighbourhood of equilibrium'.

---

1 The original prediction was that the long-run occupational wage structure in equilibrium would be explained by these variables. However, the cross section data used here will measure the state of the wage structure at a point in time only — i.e., not likely an equilibrium situation. A reasonably good R² would be predicted nevertheless since the equilibrating forces would keep the wage structure in the 'neighbourhood of equilibrium'.

---
appears to be a very good result in terms of the model and in terms of the data used for testing.

It was felt by the author, that a second way of testing the theory that was developed earlier in this paper (and at the same time testing the choice of parameters made) would be to attempt to predict the earnings in an occupation (given the number of years of schooling in the occupation) from equation 9 of Chapter II. Various sets of parameters were selected and the theoretical average earnings were calculated for the observed years of schooling in the occupations. These theoretical average earnings were then regressed against the observed average earnings. The following four regression equations typify the results.

5) \[ \bar{A}_n = 1277.1 + 0.692 T_1 \]
   \[ R^2 = 0.729 \]

6) \[ \bar{A}_n = -2236.8 + 0.614 T_1 + 162.7 P \]
   \[ R^2 = 0.819 \]

7) \[ \bar{A}_n = 1270.1 + 0.648 T_2 \]
   \[ R^2 = 0.736 \]

8) \[ \bar{A}_n = -2205.3 + 0.575 T_2 + 161.0 P \]
   \[ R^2 = 0.824 \]

The following new symbols are introduced in the above equa-
tions (the other variables are as previously defined in connection with equations 1 to 4):

T_1 this variable is calculated from equation 9 of Chapter II given the number of years of schooling in an occupation and the following parameters; r (discount rate) = 12%, L (the 'labour force life' as previously defined) = 53, A_0 (the base wage level also previously defined) = $2000, D (the cut-off point) = L - 15 years, and the slope was chosen as 35n.

T_2 is a theoretical variable calculated as above with parameters; r = 12%, L = 53, A_0 = $2400, D = n + 25, and n the slope was chosen as before as 35n.

The above four equations are representative of the type of result that can be obtained in this manner. They are not to be considered the best results that could be obtained. Time did not allow for a study of many of the possible parameter choices (since most of the parameters are of little importance one would mainly wish to explore results for various discount rates).

The explanatory power of the above equations would be expected to be no greater than in equations 1 through 4 of this chapter (and as good as in the first four equations only if the model and the specification of parameters are correct). This follows for two reasons; firstly, the regressions here are using the same independent variables as in the first four equations (schooling) only in a transformed form, secondly, it would not have been expected that a transformation of the schooling variable would explain more than n and n^2 together could explain -- this
follows from the theoretical regressions presented at the end of Part A in Chapter III in which $n$ and $n^2$ explained .999 of the total variation in the dependent variable. A comparison of the above four equations with the first four equations in this chapter confirms these expectations.

As mentioned before, however, the above regressions on theoretical variables are only to be considered representative since the possibilities open in the choice of parameters were not exhausted owing to the paucity of time available. Reconsideration of the choice of parameters since the regression work was completed suggests that a lower value for the 'base wage level', $A_0$ (in the neighbourhood of $\$1500$) combined with a lower discount rate and a less steep slope ($S_n$) would result in a theoretical variable that would be more nearly identical to the occupational earnings variable ($\overline{A_n}$) omitting random errors. That is, the constant term in an equation such as $5$ would be closer to zero and the coefficient of $T$ closer to one.
CHAPTER V

CONCLUSION AND SUMMARY

Starting with a simple concept of an occupation and a specific hypothesis about the behavior of individuals with respect to occupational choice, this paper has proceeded to build a theoretical model of the occupational earnings structure amenable to statistical testing. The underlying thesis of the paper is that the differences in occupational earnings arise because earnings as measured are gross of the returns to the investment in education which characterizes an occupation. In building the theoretical model the author became aware of the problem of the paucity of data that would allow testing of a very general model. This was especially true with respect to data on 'on the job training'. To avoid this problem, it was necessary to specify a relation between formal and informal education. Hence, both formal and informal investment in education could be measured by the one variable -- years of schooling. It would be desirable to have a model in which formal and informal education were treated strictly as two inputs in a general way, but the data made this an impossibility.

There are two types of conclusions which emerge from this type of a study. The first type of conclusion is that which follows from the most general outline of the model --
not at all dependent on the specific formulation chosen for testing but dependent on the behavioral hypotheses from which the model is built. The second type of conclusion is of a more specific nature which follows from the assumptions made in order to change the general model into a testable form.

There are two important conclusions which fall into the first category mentioned above;

1) The long-run distribution of earnings among occupations will be closely tied to the distribution of investment in human capital in the economy. The distribution of investment in human capital will, of course, be related to the demands for skill in the labour market which in turn are linked to changes in technology in the economy. It follows, therefore, that secular changes in the occupational earnings structure will be related to the occupational changes that occur in the economy -- that is, the addition of new occupations at various skill levels and the variation in training and skill requirements of the occupations that already exist in the economy. This suggests, for example, that secular comparisons of the 'skilled versus unskilled' type are incomplete without some reference to the changes in the relative degree of skill that have taken place.

2) Given that a study considers the changes in the skill distribution in the economy when changes in the earnings structure are being considered, it is also necessary to
consider changes in the discount rate or the return to investment in education. Changes in the discount rate with no changes in the skill composition would result in changes in the structure of earnings. Since this rate could change secularly (due to overall changes in the rates of return in the capital market or due to changes in the degree of imperfection in the market for funds for investment in human capital) and probably has, to discuss secular changes in the earnings structure without some reference to the discount rate would also be faulty.

With respect to the second type of conclusion -- derived from the special assumptions made about the parameters of the model and the relation between formal and informal education -- there is the result (discussed in previous chapters) that the relation between years of schooling in an occupation and the average earnings of that occupation will be a non-linear one. The prediction made in this paper was that the relation would be only slightly non-linear. This was confirmed in the empirical testing reported in Chapter IV.

The other specific conclusion that derives from this paper is that over 80% of the variation in average annual earnings among occupations can be explained by simple variables related to the time and effort invested in training for that occupation. While this is a large and significant part of the variation in occupational earnings, there still remains some 18% of the variation that is not explained by
the variables used. The immediate response is to turn to
the variables assumed unimportant in the above work -- risk,
occupational preference, and ability -- for an explanation
of the unexplained variation. A second response is to
search for better data that would give more consistent
measures of education and also that would give data for more
than one year (to allow an averaging of the earnings over a
few years to eliminate some of the short-run disturbances to
the occupational earnings structure due to demand influences.
A third response is to attempt to find a good measure of
informal training requirements for the various occupations
used in the analysis.

Future work should concentrate on these three lines of
attack, to eventually allow for construction of a model that
will explain a greater amount of the variation in occupational
incomes.
BIBLIOGRAPHY
BIBLIOGRAPHY

BOOKS


ARTICLES AND PERIODICALS


GOVERNMENT PUBLICATIONS


APPENDIX I

TABLE I --- EARNINGS-EDUCATION RELATION

<table>
<thead>
<tr>
<th>Years of Investment in Education</th>
<th>Earnings</th>
<th>First Differences</th>
<th>Second Differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$2000</td>
<td>725</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2725</td>
<td>727</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>3452</td>
<td>732</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>4184</td>
<td>742</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>4926</td>
<td>757</td>
<td>15</td>
</tr>
<tr>
<td>5</td>
<td>5683</td>
<td>780</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>6463</td>
<td>808</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>7271</td>
<td>846</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>8117</td>
<td>890</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>9007</td>
<td>947</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>9954</td>
<td>1013</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>10967</td>
<td>1094</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>12061</td>
<td>1189</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>13250</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The average Earnings (given schooling) are calculated using the following parameters; discount rate \( r \) = 12%; Retirement age = 65 \( (L = 53) \); Earnings with no education \( (A_0) = $2000 \); slope = 30n; and cut-off point = \( L - 15 \) years.
### TABLE II --- EARNINGS-EDUCATION RELATION WITH A CHANGED DISCOUNT RATE (CHANGED FROM TABLE I)

<table>
<thead>
<tr>
<th>Years of Investment in Education</th>
<th>Earnings</th>
<th>First Differences</th>
<th>Second Differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$2000</td>
<td>887</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2887</td>
<td>895</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>3782</td>
<td>909</td>
<td>14</td>
</tr>
<tr>
<td>3</td>
<td>4691</td>
<td>933</td>
<td>24</td>
</tr>
<tr>
<td>4</td>
<td>5624</td>
<td>963</td>
<td>30</td>
</tr>
<tr>
<td>5</td>
<td>6587</td>
<td>1006</td>
<td>43</td>
</tr>
<tr>
<td>6</td>
<td>7593</td>
<td>1059</td>
<td>53</td>
</tr>
<tr>
<td>7</td>
<td>8652</td>
<td>1127</td>
<td>68</td>
</tr>
<tr>
<td>8</td>
<td>9779</td>
<td>1211</td>
<td>84</td>
</tr>
<tr>
<td>9</td>
<td>10990</td>
<td>1312</td>
<td>101</td>
</tr>
<tr>
<td>10</td>
<td>12302</td>
<td>1436</td>
<td>124</td>
</tr>
<tr>
<td>11</td>
<td>13738</td>
<td>1583</td>
<td>147</td>
</tr>
<tr>
<td>12</td>
<td>15321</td>
<td>1760</td>
<td>177</td>
</tr>
<tr>
<td>13</td>
<td>17081</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. The parameters used in calculating this earnings' path are the same ones used in Table I with the exception of the discount rate which has been increased to 14%. 
TABLE III -- EARNINGS-EDUCATION RELATION WITH A CHANGED RETIREMENT AGE (CHANGED FROM TABLE I)

<table>
<thead>
<tr>
<th>Years of Investment in Education</th>
<th>Earnings</th>
<th>First Differences</th>
<th>Second Differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$2000</td>
<td>803</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>2803</td>
<td>805</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>3608</td>
<td>810</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>4418</td>
<td>821</td>
<td>11</td>
</tr>
<tr>
<td>4</td>
<td>5239</td>
<td>835</td>
<td>14</td>
</tr>
<tr>
<td>5</td>
<td>6074</td>
<td>857</td>
<td>22</td>
</tr>
<tr>
<td>6</td>
<td>6931</td>
<td>885</td>
<td>28</td>
</tr>
<tr>
<td>7</td>
<td>7816</td>
<td>921</td>
<td>36</td>
</tr>
<tr>
<td>8</td>
<td>8737</td>
<td>966</td>
<td>45</td>
</tr>
<tr>
<td>9</td>
<td>9703</td>
<td>1020</td>
<td>54</td>
</tr>
<tr>
<td>10</td>
<td>10723</td>
<td>1086</td>
<td>66</td>
</tr>
<tr>
<td>11</td>
<td>11809</td>
<td>1164</td>
<td>78</td>
</tr>
<tr>
<td>12</td>
<td>12973</td>
<td>1256</td>
<td>92</td>
</tr>
<tr>
<td>13</td>
<td>14229</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1 The parameters used in calculating this earnings' path are the same ones used in Table I with the exception of the retirement age which has been increased to 70 years.