Statistics Canada Statistique Canada

An Analysis of Earnings in Canada

By Peter Kuch and Walter Haessel

Page 152, Equation C.16 should read

$X^* = X + H = A - S - 6$

Page 157, IN9 should read

IN9 - community, business and personal services

Page 161, par. 2, line 11 should read

.... years of schooling (a_{1k}) for the British

Page 167, Equation (E.4) should read

$$EX^* = -\Sigma b_j / 2\Sigma c_j X_j = -B/2C$$

Page 175, par. 3, line 5 should read

.... or her marginal productivity,

Page 176, par. 1, line 12 should read

.... in industrial and occupational

Page 181, footnote 4, lines 2 and 3 should read

.... is $a_i OC_i + b\ (\overline{WAGE}_i)$, where a_i is the coefficient for the ith occupation's dummy, b is the coefficient for WAGE, and $\overline{WAGE}_i$ is the across-province

Page 183, Equation (F.2) should read

$$V(E) = \alpha' \Sigma_{ZZ}\, \alpha + \sigma_e^2 + \rho' \Sigma_{RR} \rho + (\sigma_H^2 + \mu_H^2)\sigma_u^2 + \beta' \Sigma_{WW}\, \beta$$

$$+ 2\alpha' \Sigma_{ZR}\, \rho + 2\, \alpha' \Sigma_{ZW}\, \beta + 2\beta' \Sigma_{WR}\, \rho ,$$

Page 197, par. 2, line 1 should read

.... withdraws from the labour market for purposes other

Page 88, footnote (3), line 2 should read

$$100\left(\frac{\partial E}{YSCH}\right) = (a_i + b\overline{SQ} + c\,\overline{KLR}),$$

Page 93, par. 2, line 2 should read
.... for the variable for overtime hours in the earnings functions for the French and North

Page 94, par. 4, line 1 should read
Within ethnic groupings, the partial ....

Page 98, par. 3, line 8 should read
in attribute levels (it can be negative). When ....

Page 102, par. 3, line 5 should read
importance by the intra-ethnic distributions ....

par. 3, line 9 should read
region per se does not contribute very much to income ....

Page 104, par. 1, line 2 should read
groups are reasonably close. However ....

Page 107, footnote (9), line 1 should read
.... is calculated as $\left[\sum_i b_i^2 S_i^2\right]^{1/2}$ where ....

Page 120, par. 2, line 7 should read
.... studies. Is it not possible that

Page 124, par. 1, line 7 should read
.... the males have higher rates-of-return on

par. 1, line 9 should read
ever-married males receive from ....

Page 125, title 7.4 should read
7.4. The Effect of the Distribution of Employment on Wages and Salaries

Page 127, footnote (2), line 3 should read

calculated as $\sum_i \bar{X}_i^f (b_i^m - b_i^f)$, ...

Page 132, Table 8.2, stub
Footnote symbol (3) should be shown on ethnic group: Other(3)

Page 147, line 1 should read
.... . If it is further assumed that ....

last par., line 6 should read
Assuming that $f(EX_i) = \theta_1 EX_i + \theta_2 EX_i^2$ in Equation 3.3 ....

Catalogue 99-758E

An Analysis of Earnings in Canada

by

Peter Kuch and Walter Haessel

CORRIGENDUM

COVER AND TITLE PAGE: THE NAMES OF THE AUTHORS SHOULD READ

By Walter Haessel and Peter Kuch

Page 6, par. 2, line 3 should read
contribution. Chapter 7 is a direct ....
bottom - names of authors should read
Walter Haessel,
Peter Kuch,

Page 33, footnote 5, line 1 should read
.... relative to others, see Chapter

Page 35, par. 1, line 8 should read
to a set of explanatory or causal factors.

Page 37, par. 1, line 3 should read
decreasing functions of the variance of the logarithm of earnings.(5)

Page 38, par. 1, line 12 should read
The intercept, α, represents ....

Page 47, footnote 3, line 3 should read
of X) is given as ....

Page 54, title 4.3 should read
4.3. Personal Attributes Variables

Page 57, par. 3, line 4 should read
.... less risk averse than persons who prefer

Page 65, par. 3, line 7 should read
.... in the province for

Page 79, Table 5.2, Col. 1 and Col. 6.
Footnote symbol (1) should be shown on headings of columns 1 and 6.

Page 82, par. 3, line 4 should read
$E_i = LWR_i + LSE_i$. Substituting ....

A HGX
904187

Census Analytical Study

An Analysis of Earnings in Canada

By Peter Kuch and Walter Haessel

Published by authority of
The Minister of Industry, Trade and Commerce

March 1979
8-0003-517

Price: $3.50

Catalogue 99-758E

Ottawa

Version française de cette publication
disponible sur demande (n° 99-758F au catalogue).

FOREWORD

The Canadian censuses constitute a rich source of information about the condition of groups and communities of Canadians, extending over many years. It has proved to be worthwhile in Canada, as in some other countries, to supplement census statistical reports with analytical monographs on a number of selected topics. The 1931 Census was the basis of several valuable monographs but, for various reasons, it was impossible to follow this precedent with a similar program until 1961. The 1961 Census monographs received good public reception, and have been cited repeatedly in numerous documents that deal with policy problems in diverse fields such as manpower, urbanization, income, the status of women, and marketing. They were also of vital importance in the evaluation and improvement of the quality and relevance of Statistics Canada social and economic data. This successful experience led to the decision to continue the program of census analytical studies. The present series of analyses is focused largely on the results of the 1971 Census.

The purpose of these studies is to provide a broad analysis of social and economic phenomena in Canada. Although the studies concentrate on the results of the 1971 Census, they are supplemented by data from several other sources. These reports are written in such a way that their main conclusions and supporting discussion can be understood by a general audience of concerned citizens and officials, who often lack the resources needed to interpret and digest the rows of numbers that appear in census statistical bulletins. For these persons, interpretive texts that bring the dry statistics to life are a vital dimension of the dissemination of data from a census. Such texts are often the only means that concerned citizens and officials have to personally perceive benefits from the national investment in the census. This particular report is one of a series planned to be published concerning a variety of aspects of Canadian life, including income, language use, farming, family composition, migration, adjustment of immigrants, human fertility, labour force participation, housing, commuting and population distribution.

I should like to express my appreciation to the universities that have made it possible for members of their staff to contribute to this program, to authors within Statistics Canada who have freely put forth extra effort outside office hours in preparing their studies, and to a number of other members of Statistics Canada staff who have given assistance. The Social Science Federation of Canada has been particularly helpful in the selection of authors for some of the studies, and in arranging for review of several manuscripts. In addition, thanks are extended to the various readers, experts in their fields, whose comments were of considerable assistance to the authors.

Although the monographs have been prepared at the request of and published by Statistics Canada, responsibility for the analyses and conclusions is that of the individual authors.

PETER G. KIRKHAM,

Chief Statistician of Canada.

PREFACE

This study constitutes an empirical examination of factors that influence the distribution of employment income received by individuals and groups of indiviuals in Canada. These factors are analyzed from two perspectives: their effect on the level of earnings of the various groups, and the influence of these factors on the variability or degree of inequality of earnings within these groups.

The distribution of economic welfare arouses a great deal of interest. In this context, the distribution of family incomes from all sources is probably a more relevant concept than the distribution of individual earned incomes. However, employment earnings form the major source of income of Canadians.

This study began as an attempt to explain the distribution of total incomes received by individuals and by families in Canada. However, a sequence of events, including the timing of data availability and changes of employment for both authors, resulted in the scope of the study being reduced to an investigation of factors that affect the level of and the degree of inequality of employment income received by individuals.

There are a number of individuals and organizations whose help and encouragement we must acknowledge. Statistics Canada and the Social Science Research Council of Canada provided financial support for research expenses and to free us from our teaching responsibilities at the University of Western Ontario. During the later stages of the study, administrators at The Pennsylvania State University were extremely tolerant as Haessel took time from his Departmental responsibilities and devoted his energies to completing an overdue manuscript.

A number of our former colleagues at the University of Western Ontario contributed to the success of this study in a variety of ways. Special mention must be made of the contributions of the following: R.A.L. Carter, who provided much valuable council on econometric problems; Lars Osberg, whose objections to the conventional human capital model lead to a reformulation of our models; and Shmuel Sharir, who read a preliminary version of the manuscript and made many useful suggestions for improving the exposition and sharpening the analysis. Four anonymous referees also provided some extremely valuable, and in some cases, incisive comments which lead to a complete rewriting of the manuscript.

Of the many research assistants who worked on this study, Richard P. Beilock of The Pennsylvania State University must be singled out for his unique contribution. Chapter 6 is a direct outgrowth of an independent project undertaken by him, and he co-authored that chapter. Capable research assistance was also provided by Heather Cohen, Alvaro Herran, Krishna Kadiyala, Linda Newton, and Jack Selody, of the University of Western Ontario. Rochelle Bush and Carolyn Grove very capably typed and retyped various parts of the manuscript.

Our greatest debt, however, is to our loving wives, Ann and Pat. During the period that this study was in progress, they had to endure much loneliness, divided attention, and occasionally sour dispositions. They bore these tribulations with understanding and usually with good humour.

Peter Kuch,
Walter Haessel,
Department of Economics,
University of Western Ontario.

TABLE OF CONTENTS

TABLE OF CONTENTS — Continued

TABLE OF CONTENTS — Continued

TABLE OF CONTENTS – Concluded

LIST OF TABLES

LIST OF CHARTS

CHAPTER 1

SCOPE OF THE STUDY

This study endeavors to explain observed differences in employment incomes received by individual Canadians. Employment incomes, or more simply earnings, consist of wages, salaries, bonuses, tips, and net income from farm and non-farm self-employment. It constitutes by far the largest component of all income received, amounting to approximately 87% of total income received by individuals. Public and private transfers and investment income are the remaining income sources and these account for approximately 8.5% and 4.5% of the total, respectively.[(1)] Thus an explanation of the distribution of employment income takes one a long way toward understanding the distribution of income from all sources.

The focus of this study is on the identification of factors that lead to differences in the size distribution of earnings received by various groups of earnings recipients. The effects of these factors are analyzed from two perspectives. The first is the effect of the factors on the level of earnings of the various groups. The second is the influence of these factors on the variability or degree of inequality of earnings among individuals within groups.

The study employs data from individual responses to 1971 Canadian Census long-form questionnaires. To a large extent it builds upon and extends Podoluk's (1968) study which employed similar data from the 1961 Census of Canada to describe selected features of the Canadian income distribution. The method of analysis used by Podoluk for the most part consisted of comparing the arithmetic mean incomes received by subgroups of the population categorized by sex, age, education, occupation, region and family status. While our study is more narrowly focused on the earnings of individuals, the method of analysis makes it possible to investigate more deeply the relationships between a variety of characteristics of individuals and the distribution of earnings among those individuals.

Many of the factors investigated in this study were considered by Podoluk. However, one important set of explanatory variables omitted by Podoluk, but which is included in this study, relates to ethnicity and immigration. Porter (1965)

See footnote(s) on page 18.

points to the importance of these variables as determinants of income in his discussion of their relationship to social class.[2] We are able to quantify many of the labor market and education phenomena he describes as being associated with ethnicity and recent immigration in terms of their effects on earnings. Some of these phenomena can be categorized as being the result of voluntary differences in behaviour among population groups whereas others may be the result of discrimination directed against certain groups.

The method of analysis used in this study is to estimate earnings functions for both males and females using multiple regression analysis. This allows the influence on earnings of many explanatory variables to be considered simultaneously. The theoretical basis of the earnings functions is the human capital model which implies that the level of earnings of an individual depends on the amount that the individual has invested in acquiring skills of value in the labor market. Thus, variables describing years and type of schooling and potential work experience are included to serve as proxies for accumulated skills and knowledge. Other variables are included which describe personal attributes (such as ethnic group, marital status, and period of immigration) which can be broadly defined as relating to labor-supply characteristics. Together with the human capital variables (representing accumulated skills and knowledge) they constitute what might be called supply-side variables. They are intended to capture those characteristics of the worker that affect his or her earnings. In addition, a set of variables representing real and potential differences in labor-demand conditions referred to as "market effects" variables is incorporated into the earnings functions. This set includes such things as occupation, industry and geographic location. The purpose of these variables is to capture those characteristics of the work situation which produce differences in the compensation paid to workers with a given set of characteristics. These estimated earnings functions are used to analyze the effects of these variables on the distribution of earnings.[3]

A summary of the study's major analytical findings is presented in Chapter 2. This discussion initially focuses on some of the major factors affecting the level of male and female earnings separately. This is followed by a comparison of the levels of male and female earnings.

See footnote(s) on page 18.

Chapter 3 presents the conceptual framework for the analysis and develops the earnings functions in a heuristic fashion. A more detailed deviation and a technical discussion of the estimation procedures are presented in separate appendices.

The data utilized in this study are discussed in Chapter 4. This includes a discussion of why the variables were selected and how they were constructed. A glossary of variable names is presented in an appendix.

The estimated male earnings equations are discussed and analyzed in Chapter 5. Groups of variables are discussed in terms of their contribution to the level of earnings and to the degree of inequality observed in earnings. A detailed discussion of the coefficient estimates is presented in an appendix. The estimated female earnings equations are interpreted in Chapter 6. The causes of intra- and inter-ethnic variations in earnings are discussed.

Chapter 7 explores the effects of sex and marital status on the distribution of earnings. Particular attention is devoted to the extent to which male-female earnings differentials are the result of sex discrimination in compensation and employment. Finally, Chapter 8 considers the distribution of actual labor supplied (in terms of weeks and hours worked) and briefly discusses the non-participants in the labor force who were omitted from the earnings analysis.

FOOTNOTES

(1) The exact allocation of the aggregate income received by individuals in the year that ended May 31, 1961 was 87.2% from employment, 8.1% from transfers and 4.7% from investment (Dominion Bureau of Statistics, Cat. No. 98-525, Bul. SX-11, Table 2). For the year ending December 31, 1970 the corresponding allocation was 86.4% from employment, 8.9% from transfers and 4.7% from investment (Rashid, 1976, Table 1).

(2) The systematic earnings differences between English and French Canadians has been the subject of a growing body of research. See, for example, Raynauld, Marion, and Beland (1964) and Boulet and Raynauld.

(3) A study published recently by Statistics Canada (Cat. No. 13-537) used a similar method to study total family income as a function of the age and education of the family head, community size, regional location, proportion of income earners in the family, etc. The objectives of that study, however, were limited to quantifying the proportion of the variance of family incomes that was explained by a variety of socio-demographic characteristics and ranking these characteristics accordingly.

CHAPTER 2

FACTORS AFFECTING THE SIZE DISTRIBUTION OF EARNINGS

2.1. The Size Distribution of Income and Earnings

There is great variability in the incomes received by Canadians. This becomes apparent from examining the 1970 distribution of income recipients across income size categories as shown in Table 2.1. Approximately 58% of income recipients received less than the mean total income of $5,033 in that year. Yet almost 4% received three or more times as much as the mean income, and 0.9% received five or more times as much as the mean income.

In 1970, employment income accounted for 86.4% of the income Canadians received from all sources. Employment income consists of wages, salaries, bonuses, tips, and net income from farm and non-farm self-employment. Investment income accounted for only 4.7% of total income. The remaining 8.9% of total income consisted of public and private transfer payments, a large proportion of which, by design, are distributed in a manner inversely related to the level of employment income. This study is concerned with explaining the distribution of employment incomes rather than total incomes.(1)

The frequency distribution of all employment income recipients by earnings-size categories (Table 2.2) exhibits much the same shape as the frequency distribution of total income size categories even though this distribution is based on slightly over two million fewer persons. The mean employment income for earnings recipients, which was approximately $5,314, is somewhat higher than the mean of income from all sources. The greatest differences between the two frequency distributions occur in the categories below $4,000. This is attributable to the concentration of government transfer payment and retirement pension recipients with no employment income in the bottom range of income size categories in Table 2.1.

Major differences exist between the earnings of males and females. This is apparent from the separate frequency distributions of male and female employment income in Table 2.2. There were 45% fewer female than male recipients of employment

---

See footnote(s) on page 33.

TABLE 2.1. The Distribution of Individuals(1) 15 Years and Over by 1970 Total Income Size Categories

Total income size group	Number of recipients	Per cent of total	Cumulative per cent
Under $1,000 (including loss)	1,874,490	16.2	16.2
$ 1,000 - $ 1,999	1,859,450	16.1	32.3
$ 2,000 - $ 3,999	2,111,965	18.2	50.5
$ 4,000 - $ 5,999	1,863,530	16.1	66.6
$ 6,000 - $ 7,999	1,569,065	13.6	80.2
$ 8,000 - $ 9,999	1,006,995	8.7	88.9
$10,000 - $14,999	895,005	7.7	96.6
$15,000 - $24,999	286,715	2.5	99.1
$25,000 and over	105,620	0.9	100
TOTAL	11,572,790	100.0	--

-- not applicable.

(1) This covers all individuals who were 15 and over on June 1, 1971 who reported receiving income in 1970.

Source: 1971 Census of Canada, unpublished data.

TABLE 2.2. Distribution of Individuals(1) 15 Years and Over by Sex and 1970 Employment Income Size Categories

Employment income size group	Number and per cent of recipients					
	Both sexes	Per cent of total	Male	Per cent of males	Female	Per cent of females
Under $1,000	1,463,415	15.5	642,155	10.6	821,260	24.7
$ 1,000 - $ 1,999	949,735	10.1	452,115	7.4	497,620	15.0
$ 2,000 - $ 3,999	1,734,625	18.4	843,205	13.9	891,420	26.8
$ 4,000 - $ 5,999	1,737,420	18.5	1,062,455	17.4	674,965	20.3
$ 6,000 - $ 7,999	1,511,775	16.1	1,224,030	20.1	287,745	8.7
$ 8,000 - $ 9,999	914,500	9.7	822,695	13.5	91,805	2.8
$10,000 - $14,999	778,385	8.3	727,065	11.9	51,320	1.5
$15,000 - $24,999	238,670	2.5	229,165	3.8	9,505	0.3
$25,000 and over	83,290	0.9	80,710	1.3	2,580	0.0
TOTAL	9,411,805		6,083,595		3,328,210	

(1) This covers all individuals 15 and over on June 1, 1971 who reported receiving employment income in 1970.

Source: 1971 Census of Canada, unpublished data.

income, 3,328,210 females compared to 6,083,595 males. The mean female employment income (approximately $3,162) was only 48% of the mean male employment income (approximately $6,574). The median earnings for females was approximately $2,775, which was an even smaller fraction (46%) of the male median of approximately $6,068. Over 50% of male employment income recipients earned at least $6,000 in 1970 compared to only 13% of female recipients.

The remainder of this chapter contains a discussion of the effects of certain factors on the level of earnings of individual males and females as well as a comparison of the earnings levels of males and to females. This discussion is based on the empirical findings of the study which are reported in greater detail in Chapters 3 through 8.

2.2. The Influence of Schooling on Earnings Levels

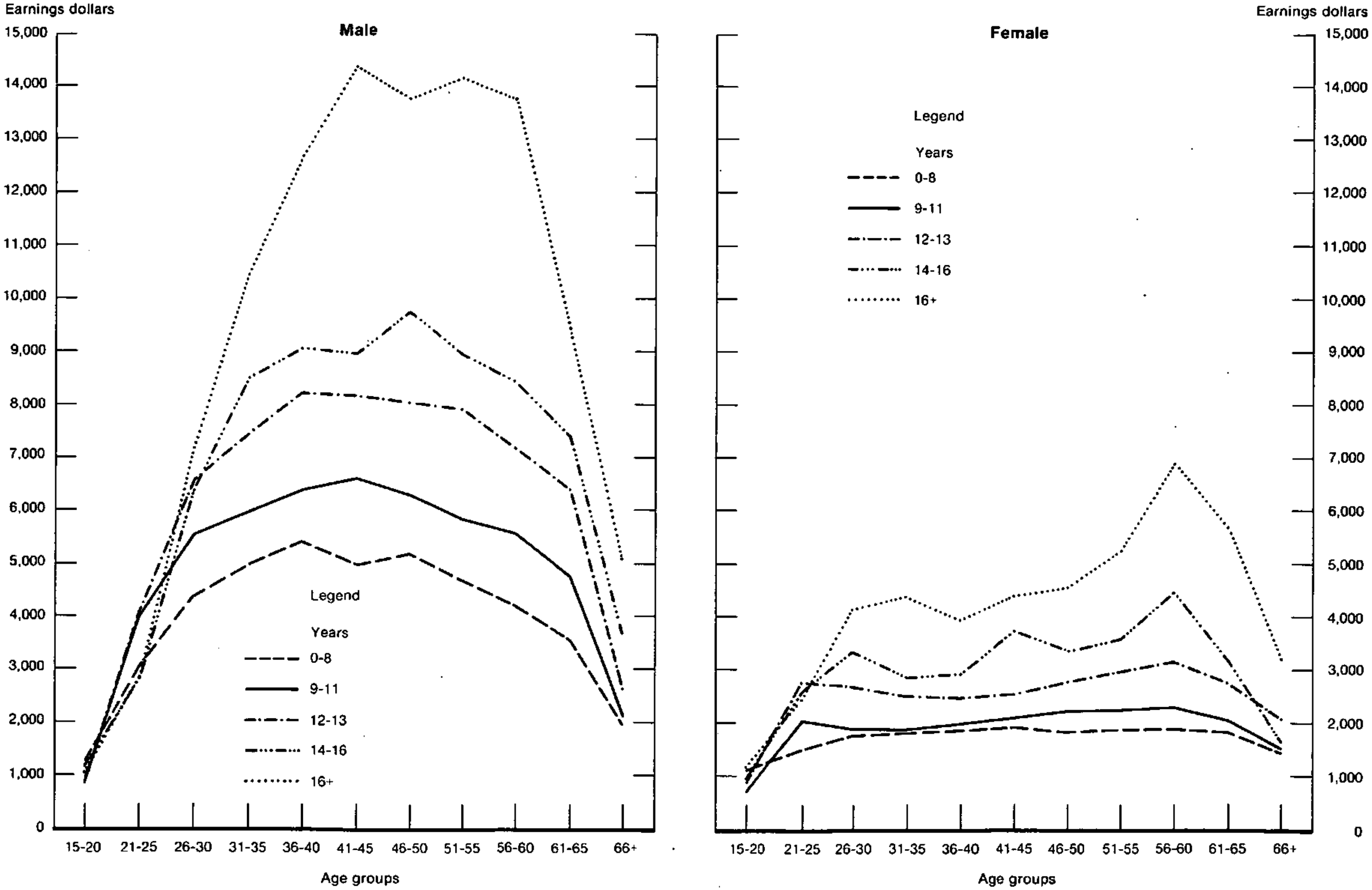
The influence of the amount of schooling on the level of earnings has been the subject of extensive investigation (this work has been summarized by Rosen, 1977). It has been well established that higher levels of schooling are associated with higher levels of earnings.

The age-earnings data for males with alternative levels of schooling summarized in Chart 2.1 are consistent with this interpretation. The geometric mean earnings of males with higher levels of education rise more steeply with age than do the earnings of individuals with less schooling. Earnings for all schooling levels peak between ages 35 and 50 and thereafter tend to fall with increasing age.

The female age-earnings data by schooling level summarized in Chart 2.1 also indicate that the higher levels of schooling are associated with higher earnings levels. The age-earnings profiles of females by education level differ qualitatively from those of the males in the sense that they do not exhibit the smooth curvature and mid-working-life peaks present for the males. For female with 13 or less years of schooling, earnings rise sharply with increasing age up to age 25 or 30 and show little change thereafter. For women with university education, earnings rise rapidly up to age 30, remain fairly constant between ages 30 and 50, and then accelerate again and peak at age 60. The age-earnings profiles of males

Chart — 2.1

Geometric Mean Earnings of Population by Sex, Age Group and Years of Education for Canada, 1970



Source: 1971 Census of Canada, Public Use Sample Tapes.

and females also differ quantitatively in the sense that the earnings peaks of the male profiles tend to be substantially higher than those of the female profiles. In fact, the peak earnings of males with 0-8 years of schooling is higher than the peak earnings of females in all education categories except those with over 16 years of schooling, and the peak earnings of male high school graduates (12-13 years schooling) are higher than those of females with over 16 years of education - (presumably university graduates).

The qualitative and quantitative differences between the male and female profiles result in part from two phenomena. First, many women, particularly married women, tend to drop out of the labor force, either permanently or temporarily, for family reasons. Thus, they tend to acquire less labor force experience than do males of comparable age. Second, even if the women do not withdraw completely from the labor force, the mothers of younger children, particularly preschool children, often only work part time. The age-earnings profiles plotted in Chart 2.1 do not take these differences into account.(2)

A second aspect of the relationship between schooling and earnings is the value of additional schooling, the so-called rate-of-return to schooling. In this study the rate-of-return to schooling is defined as the percentage increase in earnings that will result from one additional year of schooling, certain other factors held constant. Many higher paying jobs in the economy are not open to individuals with low levels of education. Hence, one of the ways that higher education levels result in higher earnings is by opening the doors to these higher paying jobs. Thus, one of the important aspects of the rate-of-return to schooling is whether this rate-of-return is computed while the occupation and industry of employment are held constant, or whether the rate-of-return includes the increased earnings associated with the "employment shift."

The rates-of-return to schooling are computed for males and females in Chapter 5 and 6. These calculations indicate that each additional year of schooling will increase earnings by 6.3% per year for males and 6.4% per year for females, when the return to schooling includes the employment shift associated with higher education. If the occupation and industry of employment are held constant, the computed rates-of-return are 5.0 and 3.5% for males and females respectively. Thus

See footnote(s) on page 33.

it appears that the employment shift aspect of education is more important for females than for males and if this is included in the rate-of-return to schooling the male and female rates-of-return are approximately equal. Within job categories, however, additional years of schooling result in smaller earnings increases for females than for males.

2.3. The Effect of Ethnic Group on Earnings Levels

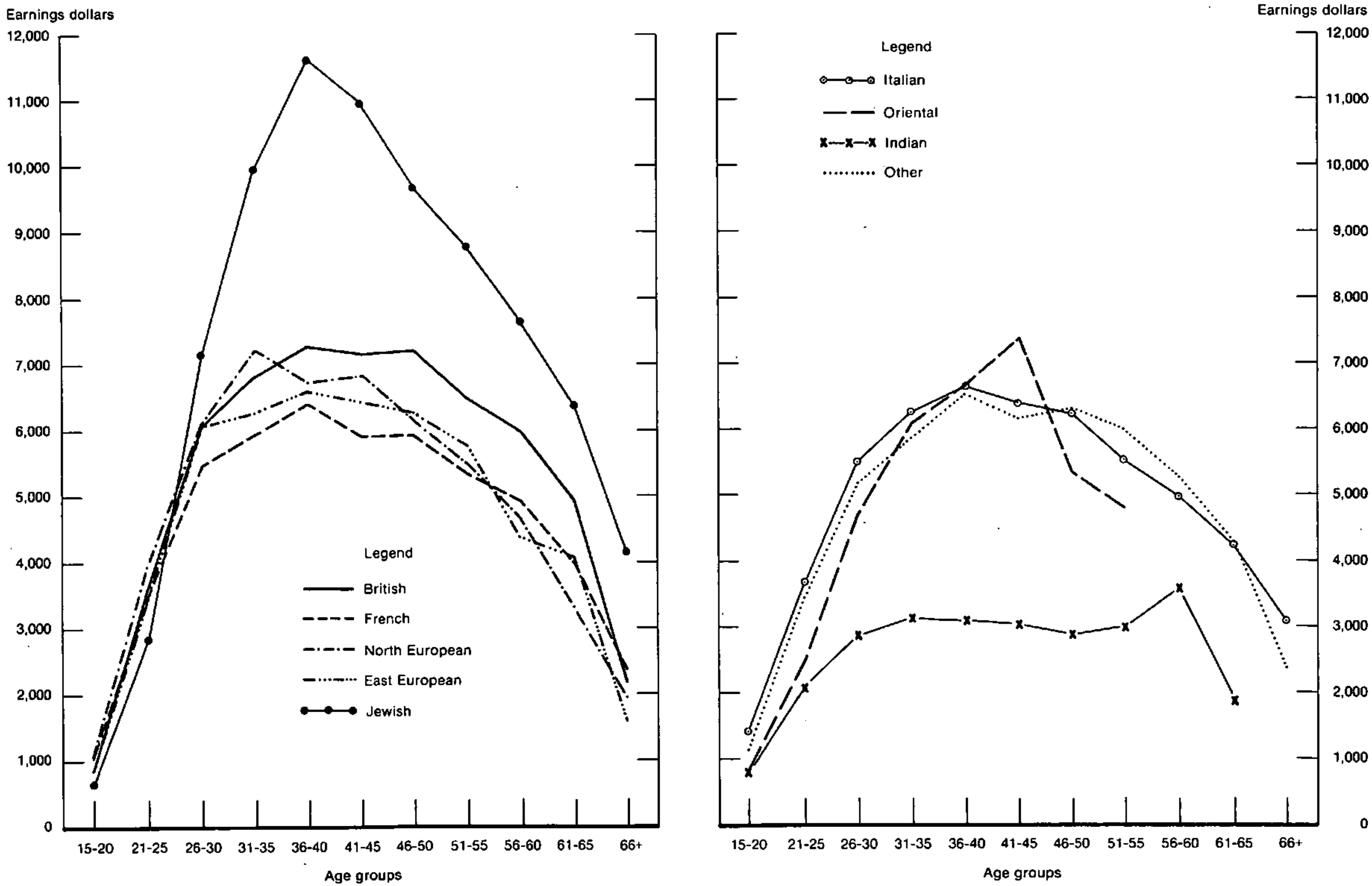
Age-earnings profiles for nine male ethnic groups are plotted in Chart 2.2. For the nine groups delineated, two differ substantially from the rest. Beyond age 30, the Jewish group has considerably higher earnings levels than any of the other groups. At the other end of the spectrum, the Native Indian group tends to have much lower earnings levels than do the other groups. The high earnings levels of the Jews can be attributed in part to the fact that this group has by far the highest level of schooling of any group (see Chapter 5). This group also has the largest proportion of labor force participants that worked full years (in excess of 48 weeks) in 1970 (see Chapter 8). The Native Indians, on the other hand, had by far the lowest level of schooling and also had the smallest proportion of full-year workers in 1970 of any of the groups delineated. Thus, motivational factors as reflected in years of school attendance and weeks worked(3) appear to be important in explaining the deviations of the age-earnings profiles of these two groups from the other groups.

The analysis for males in Chapter 5 indicates that inter-group differences in the rates-of-return to schooling are far more important in determining inter-group earnings differentials than are inter-group levels of schooling. The simple correlation between geometric mean earnings and average schooling levels across these nine groups is .59, and the simple correlation between earnings and rates-of-return to schooling including the employment shift is .74. These two effects tend to reinforce each other since those groups with high rates-of-return to schooling also tend to have high levels of schooling. The simple correlation between level of schooling and rates-of-return to schooling is .87 if the rates-of-return to schooling do not include the employment shift component, and .83 if the shift component is included.

See footnote(s) on page 33.

Chart 2.2

Geometric Mean Earnings of Males by Ethnic and Age Groups, in Canada, 1970



Source: 1971 Census of Canada, Public Use Sample Tapes.

Age-earnings profiles for five female ethnic groupings are plotted in Chart 2.3. These plots suggest that there is not very much inter-group variation in the earnings profiles. Certainly there is less variation across these groups than there was for the males. However, this is due at least in part to the fact that the smaller ethnic groupings were aggregated to produce sample sizes large enough to support the number variables being analyzed. The analysis in Chapter 6 indicates that there is also much less inter-group variation in levels of and rates-of-return to schooling for females than for males. This may also be a result of the higher level of aggregation for females.

2.4. The Effects of Sex and Marital Status on Earnings Levels

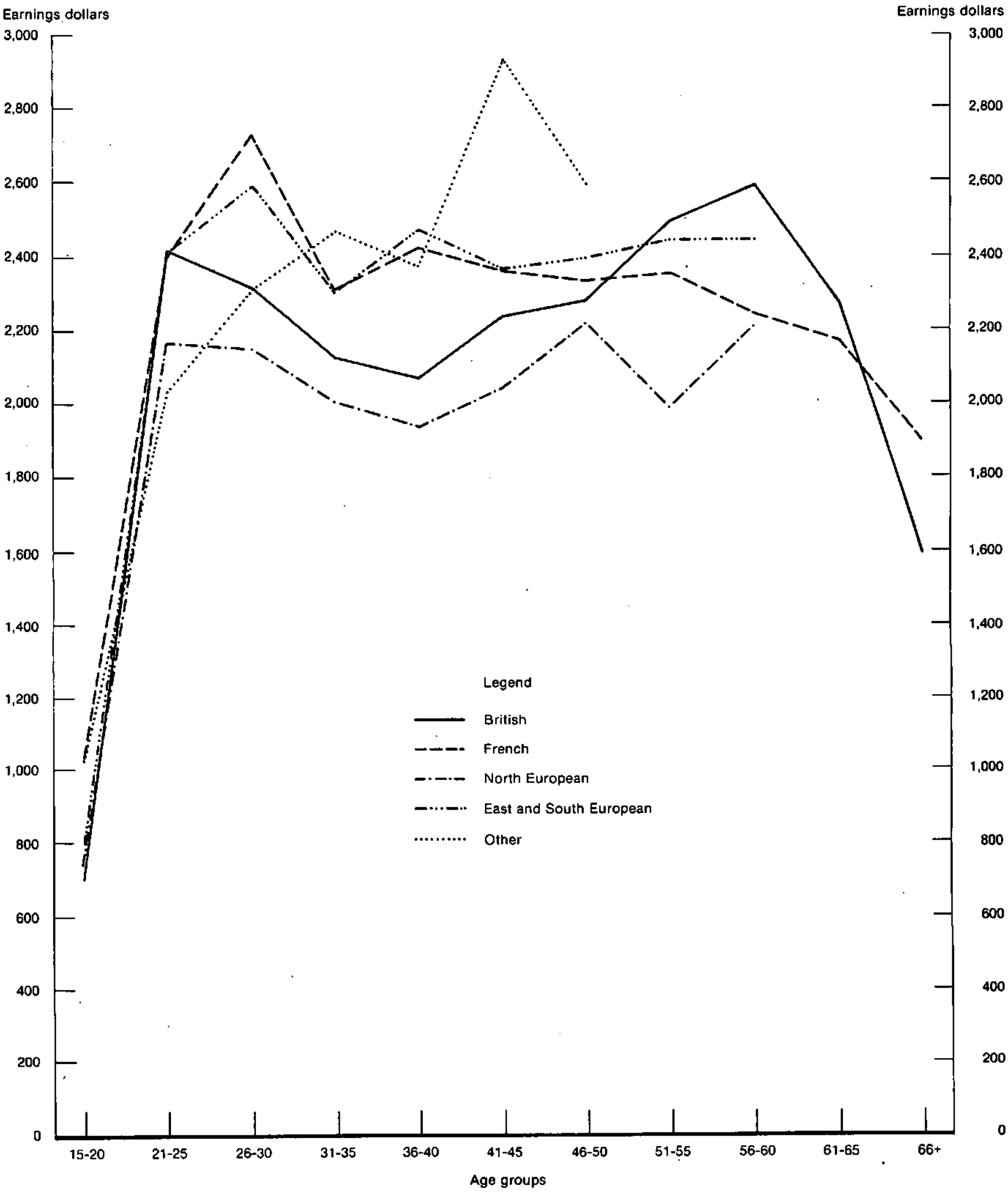
Recently there has been a good deal of interest in the level of earnings of females relative to males. The data presented in Table 2.2 indicate that females tend to have lower average earnings than males, and the data summarized in Chart 2.1 indicate this is true for all levels of education and at all age levels. This difference can at least partly be explained by the fact that many women, particularly married working women, work less than full time and this factor is not taken into account in the above distributions and figures.

A partial control for the influence of marriage on the labor-force behavior and experience of females can be obtained by separating never-married and ever-married females. Age-earnings profiles of ever-married and never-married males and females for five ethnic groups are plotted in Chart 2.4. These figures indicate that for all five groups, ever-married males have substantially higher earnings levels at every age than do all the other groups. The age-earnings profiles of never-married males and females tend to be very similar for all ethnic groups, whereas the ever-married females tend to have the lowest earnings levels. One characteristic of the age-earnings profiles of ever-married women that is common to all five ethnic groups is that earnings remain virtually constant after the women reach 25 years of age, and this characteristic distinguishes the ever-married female profiles from the rest.

Marriage has an important bearing on the number of weeks worked, years of schooling, and full-time versus part-time employment for both males and females. With respect to these characteristics, never-married women compare very favorably

Chart — 2.3

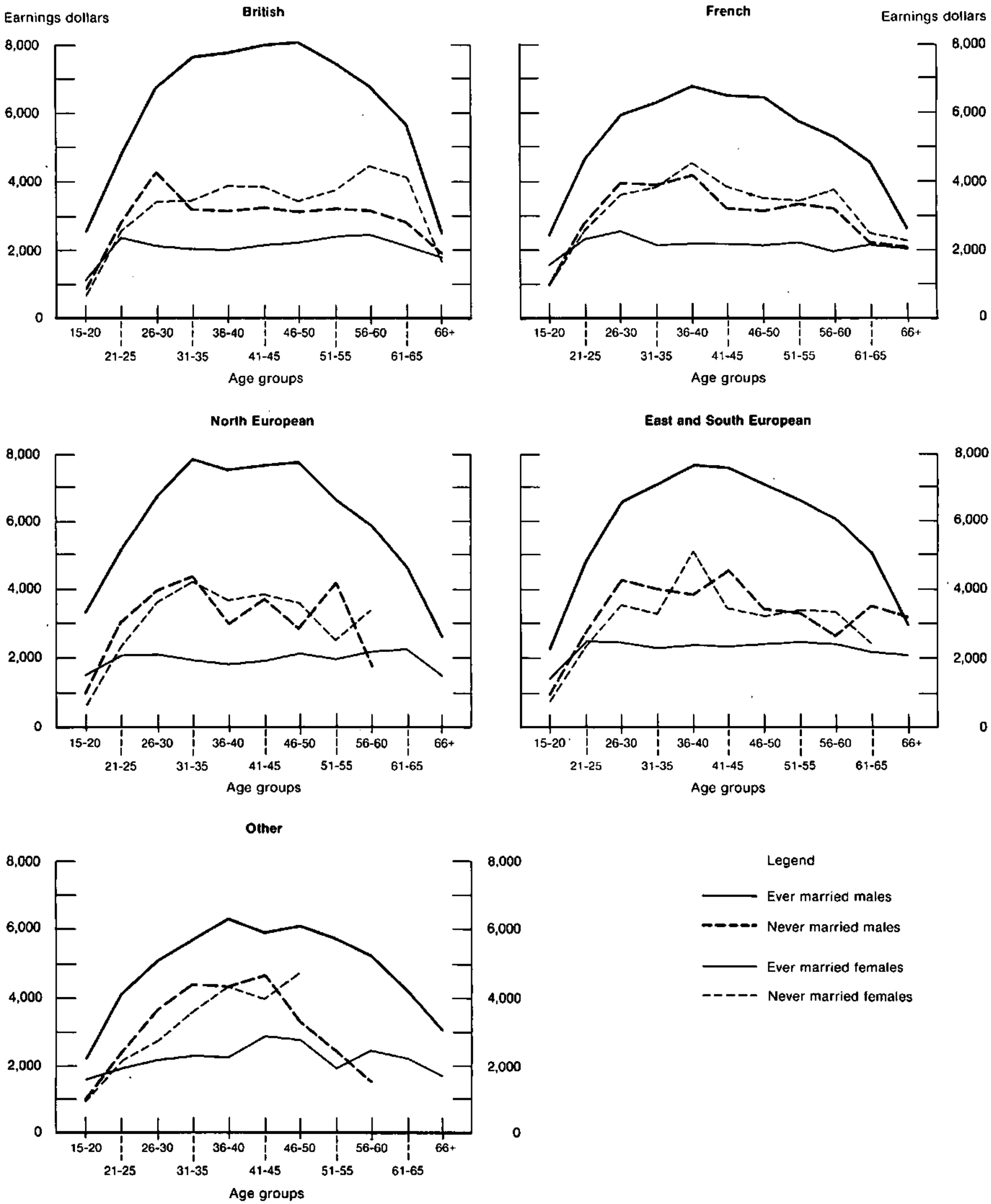
Geometric Mean Earnings of Females by Ethnic and Age Groups in Canada, 1970



Source: 1971 Census of Canada, Public Use Sample Tapes.

Chart — 2.4

Geometric Mean Earnings of Canadians by Ethnic Group, Age Group, Marital Status, and Sex, 1970



Source: 1971 Census of Canada, Public Use Sample Tapes.

with both never-married and ever-married men. Thus, an interesting question is the extent to which the earning differential between never-married females on one hand, and never-married or ever-married males on the other results from sex discrimination in the labor market.(4)

The results of calculations discussed in Chapter 7 indicate that the differences in wages and salaries between never-married and ever-married females over 30 years of age are largely the result of differences in the levels of attributes such as weeks of work, levels of schooling, geographic location, etc., and very little is the result of differences in rates-of-return to these attributes. In other words, there is no difference in the pay structure between never-married and ever-married females and the higher earnings levels of the former result from greater work effort, schooling, etc. For males, on the other hand, the situation is completely reversed. The higher earnings of ever-married males is primarily the result of differences in pay structure between the two groups with ever-married males receiving higher rewards than never-married males for given levels of work effort, schooling levels, etc. It is not clear whether this reward differential is a reflection of motivational differences between the two groups which have not been adequately controlled for, or is the result of some sort of discrimination in the labor market. The former seems to be a more plausible explanation.

In making male-female comparisons, never-married males and females have approximately equal levels of wages and salaries. This is true even though the never-married females have relatively more attractive attribute levels than do never-married males which tends to give the never-married females an earnings advantage relative to never-married males. This attribute level advantage, however, is approximately offset by the higher rates-of-return (more attractive reward or pay structures) received by the never-married males.

Comparing never-married females with ever-married males, the males have a large advantage in gross levels of wages and salaries. The two groups have roughly equal attribute levels and neither group gains any substantial pay advantage from this source. However, the ever-married males have much higher rates-of-return to their attributes than do the never-married females, and these higher rates-of-return account for the large earnings differential between these two groups.

See footnote(s) on page 33.

This large rate-of-return differential between males and females has often been attributed by other researchers to sex discrimination (see Chapter 7). However, if the rate-of-return differential between males and females is labeled as sex discrimination, the rate-of-return differentials between ever-married and never-married males should also be labeled as discrimination. However we feel the latter is probably a reflection of motivational differences between the two groups which have not been adequately controlled for in the analysis. Thus, we hesitate to argue that motivational differences between ever-married males and never-married females are unimportant, and that the wage and salary differences between males and females which are the consequences of rate-of-return differentials should be labeled as sex discrimination in pay. Labeling these rate-of-return differentials as sex discrimination would only be appropriate if the rate-of-return differentials arose solely from the employers' behavior, and such supply-related phenomena which have been omitted are unimportant.

Other researchers have argued that the principal reason for the pay of females being lower than that of males is that females are systematically excluded from higher paying jobs within broad occupational and industrial categories (Oaxaca, 1973, p. 708). In this study we have found no evidence of this effect operating in the labor market.

2.5. Sources of Earnings Inequality

The degree of inequality in the size distribution of earnings is an aspect of earnings distribution that has received considerable attention in the literature (see, for example, Atkinson, 1970). The measurement of the degree of inequality in the size distribution of earnings is somewhat arbitrary, and inequality measures are meaningful only as relative measures in comparing the degree of inequality among distributions.

The measure of earnings inequality used in this study is the standard deviation of the logarithm of earnings.(5) The size distribution of annual earnings of females is more unequal than that of males based on this criterion.

Sources of earnings inequality for males and females are discussed in Chapters 5 and 6. For females, it appears that the distribution of weeks of work

---

See footnote(s) on page 33.

is the most important source of intra-ethnic earnings inequality of all the factors considered. This is followed in order of decreasing importance by the distribution of years of estimated labor force experience, part-time work, and years of schooling. Surprisingly, region of employment was not an important determinant of female earnings inequality. This suggests that after all other factors are taken into account, region of employment per se contributes very little to the inequality of female earnings.

The investigation of sources of earnings inequality among males did not consider individual factors separately, but rater considered the following three groups of factors: personal attribute and market effects (e.g., ethnicity, marital status, region and occupation of employment, etc.), human capital effects (years of schooling, years of labor force experience, apprenticeship training, etc.), and labor-supply effects (weeks of work). The analysis shows that variations in the level of and rate-of-return to personal attribute and market effects variables are by far the largest sources of earnings inequality for males. Variations in the levels of and rates-of-return to human capital are the second most important sources of earnings inequality, and variations in weeks of work are least important. As an indication of the relative importance of the weeks-of-work effect, only 19% of the variation in annual earnings is attributed directly to variations in weeks of work among individuals, whereas 62% is attributed to variations in the weekly wage rate among individuals. The remaining 19% is attributable to the fact that those individuals with higher weekly wage rates also tend to work more weeks. In other words, there is a positive correlation between the weeks-of-work effect and the weekly wage-rate effect.

The weekly wage-rate effect is a combination of the human capital effects and personal attribute and market effects. The weeks-of-work effects are uncorrelated with the human capital effects and the positive correlation between the weeks-of-work effects and weekly wage-rate effects are entirely due to a positive correlation between weeks-of-work effects and personal attribute and market effects.

FOOTNOTES

(1) From the standpoint of economic welfare, the major concern is with the distribution of per capital family incomes rather than individual incomes. Irvine (1978) has recently examined the distribution of lifetime incomes in Canada and found that lifetime incomes are more equally distributed than are current incomes or earnings. Irvine, however, devoted little attention to the determination of current income.

(2) To a large extent, the analysis in Chapters 3 through 7 is designed to analyze these and other factors simultaneously.

(3) This argument assumes that the large proportion of Native Indians working less than 48 weeks is at least partly due to voluntary labor supply behavior and not entirely demand induced unemployment.

(4) Sex discrimination in pay has received a considerable amount of attention in the literature. This is reviewed in Chapter 7.

(5) For a discussion of the merits of this measure relative to other, see Chapter 3.

CHAPTER 3

THE CONCEPTUAL FRAMEWORK

3.1. The Characterization of the Earnings Distribution

The primary purpose of this study is to investigate the factors that lead to differences in the size distribution of earnings of various groups of individuals. In order to do this it is necessary to develop some technique for describing or characterizing entire earnings distributions such as those for males and females in Table 2.2. One way of doing this, and the procedure used in this study, is to fit some particular mathematical form to the observed frequency distributions, and then to relate the parameters of the fitted curves to a set of explanatory or casual factors.

The first step in such a procedure is to choose a mathematical form to describe the earnings distributions. Aitchison and Brown (1969, p. 108) suggest that the choice of the functional form for describing income distributions may be governed by any one or more of the following four criteria:

(a) the extent to which the form of the function is consistent with the underlying theory of earnings or income generation;
(b) the ease with which the function can be handled in analysis;
(c) the economic significance that can be attached to the estimated parameters of the distribution; and
(d) the degree to which the fitted function approximates the data.

The mathematical form chosen for this study is the density function for the lognormal probability distribution, which can simply be defined as the density function of a random variable whose logarithm obeys the normal law of probability.(1) The lognormal distribution used in this study satisfies the first three of the above four criteria. It follows logically from the underlying human capital theory of determination of individuals earnings and this theory is the basis of the earnings functions discussed in the following sections. The lognormal distributions can be entirely summarized by the mean

---

See footnote(s) on page 47.

and variance (or standard deviation) of the logarithm of the individuals' earnings. This makes the function easy to manipulate in analysis, and these parameters have straightforward economic interpretations. The mean of the logarithms is simply the logarithm of the geometric mean earnings and also corresponds to the logarithm of the median earnings.(2) Unlike the normal distribution which is symmetric, the lognormal distribution is positively skewed (i.e., it has an elongated tail to the right). As a consequence of this, the arithmetic mean earnings will be larger than the median earnings which in turn will be larger than the mode (i.e., the "peak" of the distribution). In fact, the geometric mean or median earnings level is the only pure measure of central tendency of level of earnings in the lognormal distribution.(3)

In an income or earnings distribution context, the variance of the logarithms is of particular interest as a summary measure of the degree of earnings inequality. For example, suppose it is desired to rank a series of earnings distribution according to the degree of inequality. This can be done using any one of a number of measures of inequality, one of which is the variance of the logarithms. The larger the variance of the logarithm of earnings, the more unequal the distribution of earnings. Complete equality of earnings (i.e., everyone having the same earnings) is indicated by a zero variance.(4) These distributions can also be ranked using one of several other criteria. Unfortunately, the rankings by the different criteria may not all be the same. Atkinson (1970) has shown that as long as the Lorenz curves of the earnings distributions do not cross, any measure of inequality which satisfies some simple and reasonable value judgements regarding the effect of the distribution of income on social welfare will result in the same rankings of the distributions. It is not possible for the Lorenz curves of two lognormal distributions to cross (Aitcheson and Brown, 1969, p. 113). Therefore, since the variance of the logarithms as a measure of the degree of inequality satisfies Atkinson's restrictions on a reasonable social welfare measure, it follows that most of the common inequality measures (including Atkinson's measure) result in rankings equivalent to those of the variance of the logarithms if the earnings distributions being ranked are lognormal. Indeed,

---

See footnote(s) on page 47.

it can be shown that for lognormal distributions, many of the commonly used summary measures of earnings inequality are either strictly increasing or decreasing functions of the logarithm of earnings.[(5)]

The fourth and final criterion for the selection of a mathematical form is the degree to which the function approximates the data. In a sense, this is the most important criterion since the validity of the economic interpretation of the parameters depends to some extent on how close the data are to being lognormal. This is, however, also the most difficult criterion to appraise since it should be reapplied to every different earnings or income distribution investigated. A very extensive series of tests of the hypothesis that earnings follow a lognormal distribution have been conducted by the Lyall (1968) for a large number of distributions for over 30 countries. His general conclusion can be summarized as follows: "The central part of the distribution, from perhaps the tenth to the eightieth percentile from the top is close to lognormal. But the tails of the distribution contain an excess of frequencies in comparison to the lognormal " (pp. 66-67). We have not conducted any tests to determine how well our frequency distributions are approximated by a lognormal density function because of the large number of different distributions involved. Furthermore, the advantages of using the lognormal on the basis of the first three criteria seemed to us to be sufficient to warrant its use.[(6)]

The next section contains a brief outline of the conventional human capital model and some extensions as a basis for determining the level of earnings for individuals. In the remainder of the chapter we discuss some conceptual problems encountered with using this approach.

See footnote(s) on page 47.

3.2. Theoretical Basis of the Earnings Functions

The theoretical basis of the earnings functions used in this study is the human capital model as formalized by Becker (1964) and subsequently developed by others, most notably Mincer (1974). The conventional human capital model can be written as:

$$Y_i = \alpha + r_s S_i + r_p f(EX_i) + v_i, \tag{3.1}$$

where Y_i denotes the natural logarithm of earnings of the ith individual in a particular period, S_i denotes the years of formal schooling of the ith individual, r_s is the rate-of-return to investment in formal schooling, r_p is the rate-of-return to post-school investment in human capital (on-the-job training), and v_i is a statistical error term. The amount of investment in on-the-job training is not observable and is usally considered to be a function (often quadratic) of the number of years of experience in the labour force, EX_i. The intercept, a, represents raw earning power of individuals and is the "innate ability" to earn income without either formal schooling or labor force experience.

Equation 3.1 and variations of it have been estimated using a variety of cross-sectional data bases. In this formulation, innate ability and the rates of return to schooling and on-the-job training are considered equal across all individuals. As a consequence of this, any observed variations in earnings among individuals are attributed to differences in the amount of human capital the individuals possess.

Alternatively, one could assume that rates of return vary among individuals and that different individuals have different innate abilities depending on their personal attributes. Differences in rates of return could arise either from market imperfections (discrimination and/or unequal opportunities), or from differences in ability that enable some individuals to make more efficient use of accumulated human capital. Under the assumption that human capital in time-equivalent units is homogeneous, Equation 3.1 may be rewritten as:

$$Y_i \simeq a_i + r_i H_i, \tag{3.2}$$

where r_i is the rate-of-return to investment in human capital by the ith individual, a_i is the ith individual's innate ability, and H_i is an index of the level of human capital in time-equivalent units. None of the terms on the right-hand side of Equation 3.2 are observable. Below we discuss proxies for a_i, r_i, and H_i that allow Equation 3.2 to be operationalized.

Investment in human capital requires the time of the investor as well as market resources. Indeed it is often argued that time and the associated foregone earnings are the principal inputs into the production of human capital. Thus, one can obtain a proxy for H based on time in "schooling-equivalent" units spent on the accumulation of H. We denote such a proxy as,

$$H_i = S_i + \theta f\,(EX_i), \qquad (3.3)$$

where S_i indicates time devoted to formal schooling and $f(EX_i)$ is time devoted to on-the-job training which is a function of the years of experience, and the θ is a weight to allow for the possibility that these two types of investments are substitutes in the labor market but at unequal absolute prices.

The flow of income accruing to innate ability a_i can be approximated as:

$$a_i = \alpha_0 + \sum_{j=1}^{J} \alpha_j Z_{ij} + e_i, \qquad (3.4)$$

where the Z_{ij}'s are a series of variables which control for two types of variation. The first takes account of the fact that two people with the same innate ability and human capital may, in fact, receive different income flows because of market imperfections such as discrimination or geographic segmentation of labor markets. The second type of control variable attempts to capture innate differences in such factors as motivation, aggressiveness, risk aversion, and personal discipline.[7] Note that the specification in Equation 3.4 indicates this effect has a random component e_i which varies across individuals.

See footnote(s) on page 47.

This error term is assumed to have a zero mean and constant variance.

It can be argued that the rate of return to human capital for the ith individual depends on a number of factors specific to the individual. These factors relate to either the individual's ability to make productive use of human capital in the labour market or to the availability and opportunity costs of investment funds (broadly speaking, the demand for and supply of human capital investments for the ith individual). Let this dependency be denoted as:

$$r_i = \gamma_o + \sum_{k=1}^{K} \gamma_k X_{ik} + u_i, \tag{3.5}$$

where the X_{ik}'s denote proxies for variables affecting the rate of return to human capital,[8] the γ_k's are unknown parameters, and u_i is a statistical error term with zero mean and constant variance.[9]

Thus, both the rate of return to human capital and innate ability are considered to be random variables, and both contain a systematic component (which depends on the level of certain variables) plus a random disturbance which varies from one individual to the next. By substituting Equations 3.3, 3.4, and 3.5 into Equation 3.2 and rearranging terms, the model can be rewritten as:

$$Y_i = \sum_{j=0}^{J} \alpha_j Z_{ij} + \sum_{k=0}^{K} \gamma_k X_{ik} S_i + \sum_{k=0}^{K} \lambda_k X_{ik} f(EX_i) + \varepsilon_i, \tag{3.6}$$

where $\varepsilon_i = e_i + H_i u_i$ is a combined error term, $X_{i0} = Z_{i0} = 1$ for all i, and $\lambda_k = \Theta\gamma_k$. The derivation and estimation of Equation 3.6 is discussed in detail in Appendix B.

See footnote(s) on page 47.

The usual human capital earnings Equation 3.1 is a special case of the more general model in Equation 3.6. The latter model allows for differential rates of return to schooling and on-the-job training (experience) among groups of individuals in the sample with different characteristics such as ethnic background. To illustrate the differences between these two formulations, suppose there are only two types of individuals in the sample; those having personal characteristics and attributes denoted by Z_{1j} and X_{1k}, and those having personal characteristics and attributes denoted by Z_{2j} and X_{2k}. Holding constant the level of expereince at EX*, the relationship between the dependent variable (the logarithm of earnings) and the level of schooling implied by Equation 3.6 is illustrated in Chart 3.1. The solid line represents the relationships for (say) the first group where the slope is $\sum_{k=0}^{K} \gamma_k X_{1k}$ and the intercept is $\sum_{j=0}^{J} \alpha_j Z_{1j} + \sum_{k=0}^{K} \lambda_k X_{1k} f(EX^*)$. The broken line represents the second group and the slope and intercept are $\sum_{k=0}^{K} \gamma_k X_{2k}$ and $\sum_{j=0}^{J} \alpha_j Z_{2j} + \sum_{k=0}^{K} \lambda_k X_{2k} f(EX^*)$. For the simpler model in Equation 3.1, the resulting estimates would be a weighted average of those for the two groups. A test of the hypothesis that these personal characteristics do not affect the rates of return to schooling amounts to a test of the null hypotheses that $\gamma_k = 0$ for $k = 1,2, \ldots, K$. If $\gamma_k = 0$ for $k > 0$, then γ_0 is equivalent to the rate-of-return to schooling in the usual sense associated with r_s in Equation 3.1. Similarly, a test of the hypotheses that personal attributes do not affect the level of earnings independently of their effect through the human capital variables amounts to a test of the null hypotheses that $\alpha_i = 0$, $j = 1, 2, \ldots, J$. Thus it is conceptually possible to test whether the simpler model in Equation 3.1 or the more elaborate model in Equation 3.6 is the appropriate model.

In the remainder of this chapter, we discuss some of the problems encountered with the approach outlined in this section.

Chart — 3.1

Comparison of the Estimated Earnings Schooling Relationship for Two Groups

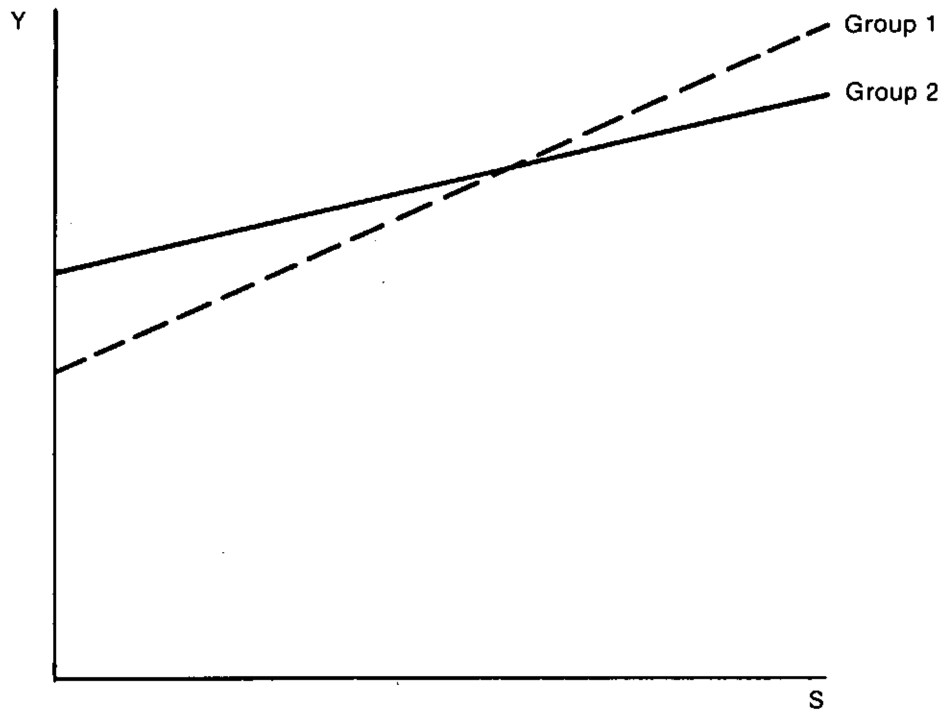
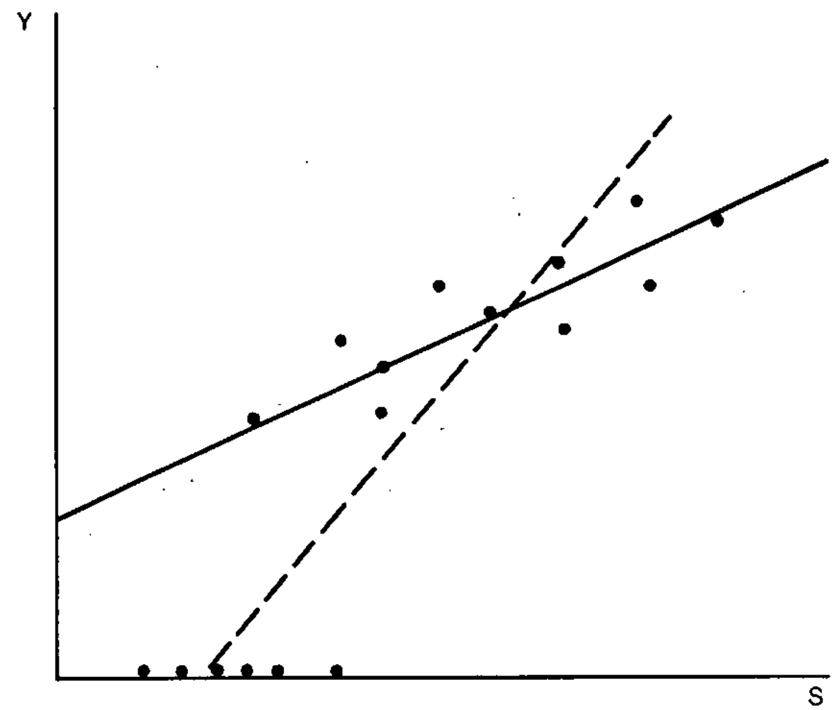


Chart — 3.2

The Problem of Estimation of Earnings Functions When Some Individuals do Not Work



3.3. Some Shortcomings of the Human Capital Model

Two conceptual problems of the human capital approach to individual income determination are briefly discussed in this section.

The first problem relates to the fact that the conventional human capital model is derived using the assumption that the individual's objective is to maximize lifetime earnings and places no value on leisure. Hence, the derived earnings equations related to potential rather than actual earnings (Blinder, 1974, pp. 14-16).(10) This raises a question regarding the appropriate specification of the dependent variable in cases where the individual works other than full-time or full-year. If the logarithm of annual earnings is the dependent variable and the individual voluntarily chooses to work only part-year or part-time, then some control for the amount of labor supply should be included. Annual earnings equal the number of weeks worked multiplied by the average weekly wage rate. Taking logarithms results in E = LWW + LWR, where E, LWW, and LWR are respectively the logarithm of annual earnings, weeks worked, and the weekly wage rate. Two specifications for the dependent variable are possible; the logarithm of annual earnings or the logarithm of the weekly wage rate. We report results for both specifications, and when E was used as the dependent variable, LWW was included as a control variable.(11)

The second conceptual problem relates to the fact that in the traditional derivation of the human capital earnings equation for an individual, it is assumed that the demand for the individual's labor services is infinitely elastic. While this may be an appropriate assumption for the individual, Osberg (1975) and Tinbergen (1975) have noted that the distribution of earnings of all individuals in society will depend on the demand for the particular types of labor relative to their supply. For simplicity, assume there are only two types of labor, A and B. All workers are identified as being one of these two types on the basis of their education, experience, etc. At any point in time, the derived demand for labor type A in a region will depend on the industrial structure of the region, the demand for the products of the various industries or firms in the region, the level of capital, and the wage rates for both labor types. Similar considerations apply to labor type B. The supply of the two types of labor in the region depend on the distribution of the level of human capital among individuals and the two wage rates.

See footnote(s) on page 47.

Equilibrium will require that the wage rates adjust so that the demand is equal to the supply for each of the two labor types. Thus, the pattern of wage rates in a region will depend on aggregate demand characteristics of the region as well as the labor supply characteristics.(12) Consequently, some additional variables, subsequently referred to as market effects variables, were included in the specifications to serve as controls for variations of the occupational and industrial mix of the economy among regions. The specific market effects variables included in the model are discussed in Chapter 4, but they include items such as the provincial capital/labor ratio in the industry of employment and variables to control for region, type of community, and industry and occupation of employment. The expanded version of the model including the market effects variables is subsequently referred to as the full model, and the specification without them is referred to as the reduced model.

3.4. The Measurement of Labor Force Experience

In the human capital model, post-school investments in human capital through on-the-job training are usually considered to be a function of the years of labor force experience. Unfortunately, very few cross-sectional data bases of the type used in this study include information on this crucial variable. The standard procedure is to define potential labor-force experience for the individual as the individual's age at the time of the survey minus the individual's age after completion of school, where the latter quantity is approximated as years of schooling plus six. This potential experience measure is then used as a proxy for actual experience.

The crucial assumption involved in this procedure is that the individual has worked continuously since leaving school. The assumption may be a reasonable one for males since they tend to have a fairly strong commitment to the labor force. Females, however, have traditionally had a much more tenuous attachment to the labor force.(13) In particular, they tend to leave the labor force after an initial period of

See footnote(s) on page 47.

post-schooling employment to give birth and rear children. They may resume labor market activity between children once the last-born is old enough to be cared for by someone else, if family circumstances so require. More frequently, if they return to work at all, they do so on a continuous basis at some junction after their final child is regularly attending school or has left home (Mincer and Polacheck, 1974). During these hiatuses in work activity they tend to cease investing in human capital which is rewarded by the labor market, and such human capital as they acquired prior to dropping out of the labor force tends to depreciate. Thus, for females, the measure of potential work experience used for males (age minus years of formal schooling minus six) is an unsatisfactory proxy for actual work experience. Consequently, we have devised a method which can be used to estimate labor force experience adjustments based on the number of children a woman has had. The procedure involves two stages. The first stage involves the estimation of an earnings function of the type outlined above but which includes two additional variables: the number of children ever born and the number of children ever born squared. From these first stage estimates it is possible to determine the effect on predicted female earnings of having given birth to any number of children. It is then possible to solve for the increased amount of labor force experience required to offset the loss in earnings due to childbearing. This is interpreted as the "experience equivalent" of the number of children born to a woman. This experience equivalent is then used to adjust potential experience to obtain a revised estimate of actual experience, which is age minus years of formal schooling minus six minus the experience equivalent. These revised estimates are then used to re-estimate the earnings functions. This procedure is discussed in detail in Appendix C.

3.5. The Problem of Nonworking Individuals

The samples used to estimate the models in this study have been restricted to those individuals who actually worked in 1970 and consequently had positive earnings. This procedure was dictated to some extent because the dependent variable in the earnings functions was the logarithm of annual (or weekly) earnings, and the logarithm of zero is undefined. Restricting the sample to individuals who worked, however, results in what has become known as sample selectivity bias (Heckman, 1976b; Wales and Woodland, 1977).

The problem is that by restricting attention to only those individuals who work, there will be a tendency to underestimate the return to economic attributes such as schooling. This is illustrated in Chart 3.2., where a number of hypothetical observations of earnings have been plotted against the level of schooling.(14) The solid line represents an ordinary least-squares estimate of the relationship between earnings and years of schooling when the sample has been restricted to the individuals with positive earnings, and the broken line represents an estimate based on the entire sample. It is clear that the estimated "return to schooling" (slope) in the latter case is much higher than in the former. It can be shown that using ordinary least squares with the entire sample gives excessive weight to the non-working individuals and will result in an over estimate, whereas the restricted sample results in an under estimate of the return to schooling. Clearly the problem is more severe for females than for males since a much higher proportion of females do not work. Heckman (1976b) and Wales and Woodland (1977) discuss a number of estimation procedures that are appropriate. Unfortunately we were not aware of this problem until this study was virtually completed.(15)

The specific data used in the study are discussed in the next chapter.

See footnote(s) on page 47.

FOOTNOTES

(1) More specifically, if X is a random variable restricted to positive values ($0 < X < \infty$) and $Y = \ell nX$ is a normally distributed random variable, then X is said to be lognormally distributed, and the density function can be denoted as:

$$f(X) = (2\pi\sigma^2X^2)^{-\frac{1}{2}} \exp\{-(\ell nX - \mu)^2/2\sigma^2\}$$

(2) Let $X_1, X_2, \ldots, X_n$ represent the earnings of the n recipients in the group. The mean of the logarithms is defined as $\bar{Y} = \sum_{i=1}^{n} \ell n(X_i)/n$, where ℓn denotes the natural logarithm. The geometric mean is defined as the antilog of $\bar{Y}$, or $G = \exp\{\sum_{i=1}^{n} \ell n(X_i)/n\}$. For the lognormal distribution, G corresponds to the median level of earnings which is defined as that earnings level such that exactly 50% of the earnings recipients receive less than the median earnings and 50% receive more than the median. In other words, it is the earnings level in the "middle" of the distribution. It should be noted that unless all earnings are exactly equal, G will always be less than the arithmetic mean which is defined as $\bar{X} = \sum_{i=1}^{n} X_i/n$. The reason for this is that the geometric mean gives relatively more weight to low incomes than to high incomes.

(3) Let $Y = \ell nX$, where Y is a normal random variable with population mean μ_y and variance σ_y^2. Then the population mean of X (i.e., the expected value of X is given as $\mu_x = \exp\{\mu_y + \frac{1}{2}\sigma_y^2\}$, while the population median is $M_e = \exp\{\mu_y\}$, and the population mode is $M_o = \exp\{\mu_y - \sigma_y^2\}$. Thus, both the arithmetic mean (which is an estimator of μ_x) and the mode of X depend on the degree of dispersion (σ_y^2) as well as the level of Y. Interested readers are referred to Aitcheson and Brown (1969, Ch. 2) for more details.

(4) Intuitively, this can be seen from the formula for calculating the variance which is defined as $S_Y^2 = \sum_{i=1}^{n} (\ell nX_i - \bar{Y})2/(n-1)$, where $\bar{Y}$ is the mean of the logarithms of income. If everyone has the same earnings i.e., $X_1 = X_2 = \ldots = X_n$, then $\ell nX_i = \bar{Y}$ for all i and $S_Y^2 = 0$. If everyone has incomes that are relatively close to the mean then $(\ell nX_i - \bar{Y})$ will tend to be small for all individuals and S_Y^2 will tend to be small. However, if some people have very large earnings and others have very small earnings relative to the geometric mean then some of the terms in the numerator will be large and S_Y^2 will tend to be large.

(5) Following Aitchison and Brown (1969, Ch. 2 and 11), and using the same notation as in the preceding footnotes, the population variance of X is given as $\sigma_x^2 = (\exp\{\sigma_y^2\} - 1) \cdot \exp\{2\mu_y + \sigma_y^2\} = \mu_y^2\eta$, where $\eta = \exp\{\sigma_y^2\} - 1$ is the square of the coefficient of variation. Note that the coefficient of variation, a commonly used measure of inequality, is an increasing function of the variance of the logarithm and does not depend on the mean. Similarly, the Gini concentration ratio (the area between the Lorenz curve and the diagonal of complete equality) is an increasing function of the variance of the logarithms and is given as $C = 2F(\sigma_y/\sqrt{2}) - 1$, where $F(t) = \int_{-\infty}^{t} (2\pi)^{-\frac{1}{2}} \exp\{-Z^2/2\}\, dZ$ is the cumulative probability distribution of a standard random variable.

(6) Other distributions could be used of course. Metcalf (1972) used a three-parameter distribution called a displaced lognormal, while Thurow (1970) used the beta distribution to characterize income distributions in the United States.

(7) The specific variables included in Z_{ij}'s are discussed in Chapter 4.

(8) The variables included in the X_{ik}'s and the justification for them are discussed in Chapter 4.

(9) The derivation of Equation 3.5 is discussed in detail in Appendix A.

(10) Some recent attempts have been made to investigate the consequences of relaxing this assumption. See Blinder and Weiss (1974), Ghez and Becker (1975), and Heckman (1976a).

(11) If the coefficient of LWW is restricted to be equal to one, the specification with E as the dependent variable is equivalent to the specification with LWR as the dependent variable. Inclusion of LWW as a control variable can result in a problem of simultaneity bias since weeks worked and annual earnings are simultaneously determined. However, using LWR as the dependent variable will result in a specification error if the coefficient of LWW is different from one. Conceptually, the solution would be to estimate a system of equations which simultaneously explain weeks worked and the wage rate. We have not attempted this refinement because of the difficulty involved in doing it with the data set used in this study (see Chapter 4).

(12) Symbolically the demand for labor of type k can be denoted as

$$D^k = G^k(W^A, W^B, [K_{ij}], [P_{ij}]),$$

where W^A and W^B denote the wage rates, $[K_{ij}]$ is a vector with elements representing the capital stocks of firm j in industry i, and $[P_{ij}]$ is a vector of product prices reflecting final demand conditions. The supply of labor of type k can be denoted as

$$S^k = J^k\ (W^A,\ W^B\ [H]),$$

where [H] is a vector of human capital characteristics and other personal attributes of the individuals in the region. Equilibrium requires that $S^k = D^k$ for k = A, B. Thus, the wage rates are simultaneously determined. Solving for the wage rates as functions of the exogenous variables results in reduced form equations that include human capital and other attribute variables from the supply side as well as indicators of the level of capitalization and industrial structure of the region from the demand side.

(13) For example, see the evidence on the median number of years in the labor force by sex for different age-groups reported by Statistics Canada, (Catalogue No. 13-557, Table C).

(14) Actual earnings have been plotted since the logarithm of zero earnings for nonworkers is not defined.

(15) We are grateful to an anonymous reviewer of an earlier draft of the study for drawing our attention to this problem.

CHAPTER 4

THE DATA AND THE VARIABLES

This study is based primarily on data gathered from responses to the census long-form questionnaire administered in June 1971. These data were extracted from the Individual File of the 1971 Census Public Use Sample which represents a 1% sample of individual Canadians. The analyses are based on the subsets of males and females who satisfied the following criteria:

(a) were older than 14 years in June 1971;
(b) were either native-born Canadians, or had immigrated to Canada prior to 1971;
(c) were not living in Prince Edward Island, the Yukon or the Northwest Territories in June 1971;
(d) were not unpaid family workers during the last week of May, 1971 (if unemployed during that week, this status is determined by the job of longest duration held since January 1, 1970);
(e) had worked and had positive employment income in 1970;
(f) did not have either wages and salaries or self-employment income in excess of $74,999 in 1970 (if both female and living in the Maritime Provinces, the limit was $49,999); and
(g) had reported both an industry of employment and an occupation which could be identified.

The resulting sample sizes were 54,168 males and 31,481 females.[1]

Data on individual characteristics were obtained from the Public Use Sample (hereafter referred to as the PUS). Other Statistics Canada data, both published and unpublished, were used to construct variables descriptive of work and schooling environments that were not specific to individuals.

All of the variables employed in the analyses were selected on the basis of the theoretical model presented in the preceding chapter. They either directly represent one of the theoretical constructs developed there, or are indirect or

---

See footnote(s) on page 68.

partial proxies for such a construct. With all of its richness of variables, the PUS as a data base has a number of weaknesses which limit a researcher's ability to explain the variation in individual incomes. They affect both the choice of the analytical method used and the accuracy with which the actual data variables represent the theoretical variables in the model. In this regard, the two most prominent weaknesses of the PUS are first, the categorical nature of certain key variables; and second, the questionable correspondence between the work activity producing the earnings reported and the work activity described by the data. Concerning the former, the gross categories in which weeks worked and hours-usually-worked per week are presented prevent the construction of wage-rate and labor-supplied variables accurate enough to warrant the estimation of simultaneous-equation models. Concerning the latter, the information on individuals' occupation, industry, class of worker, and hours-usually-worked-per-week for the most part relate to the jobs held in the last week of May, 1971,(2) whereas the annual earnings reported related to 1970. The job held in May, 1971, may not be the same job that produced the bulk of the earnings in 1970, thereby introducing significant measurement error into the analysis.

Due to the nature of the data, many of the variables used in the analysis are indicator or dummy variables as opposed to continuous variables. Dummy variables are assigned the value of unity if the individual belongs to a particular category, and zero otherwise. For example, a set of dummy variables was defined to account for the five regions of the country. Thus, if the individual was living in the Maritimes, the maritime dummy was coded one and the dummies for the other regions were coded zero for that individual. Where a series of dummy variables form an exhaustive set of categories, one member of the set must be excluded from the equation for mathematical reasons.(3) Thus, only four dummy variables are required to describe the five geographic regions. The coefficient of a particular dummy variable is interpreted as the difference in the intercept for that category relative to the excluded category. Continuing with the region example, dummy variables were coded for the Maritimes, Quebec, Ontario and the Prairies; British Columbia was the excluded region. Thus, the estimated coefficient for the (say) Quebec dummy is interpreted as the predicted

---

See footnote(s) on page 68.

difference in the dependent variable between Quebec and British Columbia when all other factors are held constant. Since it is always possible to derive the set of coefficients that would have resulted from excluding the dummy for a different category by simple arithmetic, the choice of which category to exclude is arbitrary and this choice is normally made on the basis of expository convenience.

The remainder of this chapter is concerned with a discussion of the basic variables used in the analyses. For ready reference all the basic variables used in the study are listed in Appendix D with identifying notations and short definitions. Interaction variables which are formed by multiplying basic variables together arennot discussed.

4.1. Dependent Variables

Two dependent variables are used in the analyses throughout most of this study. The first, E, is the natural logarithm of annual earned income - l.e., income from wages and salaries plus income from self-employment. Income from wages and salaries is gross (exclusive of deductions) and includes military pay and allowances, tips, commissions and bonuses, and piece-rate reimbursement. Income from self-employment includes farm income (net of depreciation and operational expenses) and income from business or professional practice (net of operational expenses). The second dependent variable (LWR) is the natural logarithm of the implicit weekly-wage rate. The implicit weekly-wage rate is the quotient of annual earned income and the number of weeks worked in 1970. Since the weeks worked appears in the PUS in a categorical form (0, 1-13, 14-26, 27-39, 40-48, and 49-52 weeks), it was first converted to a continuous variable by assigning to each category its midpoint value. Individuals who had not worked at all in 1970 (i.e., for whom weeks worked was coded as zero) were excluded from the sample.

An intersex comparison of earnings is undertaken in Chapter 7. Self-employment earnings are excluded from the dependent variable in that analysis because one of the principal questions investigated is the extent of sex discrimination in wages and salaries. Hence, income from self-employment was excluded because self-employment earnings presumably would not reflect sex discrimination in the same way as wage and salary earnings.

4.2. Independent Variables

For expository convenience, all of the basic explanatory variables have been grouped under three separate headings - personal attributes, market effects, and human-capital characteristics. It is recognized that this taxonomy may not be entirely defensible on theoretical grounds. Where the appropriateness of the group assignment of a particular variable is open to question, the reader is asked to accept it as a matter of convenience.

The personal attributes and market effects variables correspond to the Z variables referred to in Chapter 3. In addition, some of these variables affect the accumulation of human capital or its rate of return, consequently they also serve as X variables. Such joint use is indicated where the particular variable is discussed.

4.3. Personal Attributes Variable

4.3.1. Ethnic Group

The set of ethnic groups used in the analysis of male earnings differs from that used for females. For the males, nine dummy variables were defined for 10 groups from the Ethnic or Cultural Group Field so as to represent the ethnic origin groupings most likely to explain differences in individual earnings. The concern here is to capture major differences in culturally imbued labor-market behaviour and foci of labor-market discrimination, and not simply to list the numerically most important Canadian ethnic groups. The categories employed in the male analysis are individuals claiming the following descent: French (ETH1); Negro or West Indian (ETH2); North Europeans - Austrian, Finnish, German, Dutch or other Scandinavian (ETH3); East Europeans - Czech, Hungarian, Polish, Russian, Slovak, or Ukranian (ETH4); Italian (ETH5); Jewish (ETH6); Oriental - Chinese or Japanese (ETH7); Native Indian (ETH8); Other and Unknown (ETH9). British Isles ancestry is the excluded category.

For females, these 10 categories were aggregated into five groups: British, French and North European were left unchanged. The East European, Italian, and Jewish groups were combined to form a group called East and South European.

The Negro-West Indian, Oriental, Native Indian and Other and Unknown groups were combined into a group called Other.

Besides having a direct effect on earnings, thus being among the Z's referred to in Chapter 3, these Ethnic dummies are assumed to indirectly affect earnings by affecting the rates of return to schooling and experience. As such, they are also included among the X's.

4.3.2. Marital Status

It seems reasonable to expect that being married and thus having a family, as opposed to being single, motivates a male to increase his earnings. For a female, marriage may reduce the need for earnings relative to that of a single female. On the other hand, it is not clear whether being widowed, divorced or separated is likely to increase or decrease the need for earnings because of the possible presence of children on one hand, and insurance or alimony payments on the other. In order to capture these motivational effects associated with marital status, the PUS Marital Status codes were grouped into two dummy variables - currently married (MS1), and currently widowed, divorced or separated (MS2), leaving never-married as the excluded category. An ever-married category (EM) is also used in connection with females.

Being the head of a family unit and having the economic responsibilities associated with it, may not be uniquely defined by the intersection of sex and martial status. Widowed, divorced or separated persons may or may not have unmarried children living with them. Single men or women may have their natural children, or a guardianship child or ward under 21 years-of-age living with them. In order to capture the motivation for earnings associated with these sorts of dependency situations, a dummy variable (HH) that takes the value of unity if an individual is the head of a census family, and zero otherwise, was also used in the analyses.

The timing of marriage is likely to have a significant effect on educational attainment and consequently, on observed earnings. Particularly for a male, getting married is likely to increase the opportunity cost (or disutility) of time devoted to education.[4] Married individuals would continue formal schooling after marriage only if it would yield a relatively high rate-of-return. Therefore, we might expect those who married prior to the completion of their formal schooling to have higher rates-of-return to schooling than individuals who married after completion of schooling. In order to capture this phenomenon, a dummy variable (MIS) was added to the analysis which takes the value of unity if the individual's age at first marriage (taken from the PUS field of that name) was less than or equal to the years of schooling attended plus six. MIS is clearly one of the variables belonging in the X group, but since it might serve as a proxy for attributes having a direct effect on earnings it is included among the Z's as well.

4.3.3. Bilingual Ability

The ability to speak both official languages (English and French), as opposed to only one of them, might well be expected to lead to higher earnings. For the French Canadian historically, a knowledge of English has been required for access to higher paying jobs. More recently, the increased insistence by francophone Canadians on being served in their own language and a public policy making the the Civil Service effectively bilingual, has made job advancement of anglophones in a number of settings contingent on having a working knowledge of French. Furthermore, the acquisition of additional language skills in order to make oneself more marketable can be viewed as being associated with motivational characteristics that ultimately lead to higher earnings. In order to capture phenomena such as these, a dummy variable (LNO2) that is coded one if an individual is able to speak both English and French, and zero otherwise, was added to the analysis.

4.3.4. Migration

Migration can be viewed as a human capital investment (Sjaastad 1962). However, it can also be viewed in a more psychological and sociological context. Individuals who pull up stakes and move to another country or to another part of the same country can be presumed to be more achievement oriented, and more inclined to accept risk and uncertainty than individuals who do not move. Achievement

---

See footnote(s) on page 68.

orientation and willingness to accept risk, all other things equal, should lead to higher earnings. However, the newly arrived immigrant is likely to face linguistic and other cultural obstacles that tend to depress earnings. Over time, it can be expected that the immigrant will become increasingly assimilated into Canadian society, resulting in an amelioration of these difficulties and a consequent rise in earnings (Porter, 1965, Chapter 2).

Five dummy variables have been used in the analysis in an attempt to capture these conflicting effects of immigration on earnings. One variable, MIG, is assigned a unitary value if a person was living outside of Canada in 1966. Another focuses on internal migration: IMIG is assigned a unitary value if the individual was living in a different province or county within Canada in 1966. Both MIG and IMIG are constructed from the PUS Residence on June 1, 1966 Field. Another set of three variables, constructed from the Period of Immigration Field, is used to differentiate between Canadian- and foreign-born individuals and to indicate when immigration took place. If an individual immigrated to Canada before 1946 the variable PIM1 is set equal to unity, if the immigration took place between 1946 and 1960 the variable PIM2 is set equal to unity, and if it took place between 1961 and 1970 the variable PIM3 is set equal to unity. Being Canadian-born is the excluded category.

4.3.5. Self-Employment

Opportunities aside, the personality traits that lead a person to prefer self-employment to being an employee may tend to produce systematically higher or lower earnings. A preference for being one's own boss could stem from being more enterprising, assertive, innovative or less risk aversive than persons who prefer to work for someone else. These traits are normally associated with higher earnings. Alternatively, individuals may choose self-employment because of the freedom of choice regarding hours of leisure associated with self-employment, which in turn might be expected to produce lower earnings. Either line of reasoning suggests the inclusion of a self-employment variable (SE) in the explanation of earnings. SE is constructed directly from the PUS Class of Worker Field and is assigned the value of unity if the individual owns the business (incorporated or unincorporated) in which he or she is employed, and it is assigned the value of zero if the person works solely for wages, salary, tips or commission. The personality traits associated with self-employment can be presumed to operate on earnings both directly, putting it in the class of Z variables, and indirectly through the rates-of-return to human capital, making it a candidate for the class of X variables as well.

4.3.6. Labor Supplied

Annual earnings can be thought of as the product of a weekly-earnings rate and the number of weeks worked per annum, and, in turn, the weekly-earnings rate can be viewed as the product of an hourly-earnings rate and the number of hours generally worked per week. Thus, variations in the hours generally worked per week and in the weeks worked per annum as well as variations in earnings rates will result in corresponding variations in annual earnings. Since economic theory predicts that hours and weeks of labor supplied will depend, among other things, on the corresponding hourly and weekly wage rates, a complete explanation of annual earnings should involve the simultaneous explanation of the amount of labor supplied and the wage rate. Unfortunately, the form of the PUS data does not allow using simultaneous estimation procedures. As an expedient to deal with this labor supply problem, dummy variables for the hours normally worked per week are introduced into both the weekly- and annual-earnings equations, and a continuous variable controlling for weeks worked is introduced into the annual earnings equation.

Concerning the hours normally worked per week, 35 to 44 hours is treated as a norm: for persons working fewer hours a part-time-hours variable (PTHRS) is coded one, and for persons working more than 44 hours an overtime hours variable (OTHRS) is coded one. Since the Hours Usually Worked Field upon which these variables are based refers to the week prior to enumeration (or for those unemployed during that week, the job of longest duration held since January 1, 1970), experimentation was performed with an alternative part-time weeks variable (TM1) based on the 1970 Work Activity Field.[5]

The weeks worked continuous variable (LWW) is constructed from the categories of the PUS Weeks Worked During 1970 Field in the following manner:

See footnote(s) on page 68.

if 1-13 weeks, LWW = ℓn 7,
if 14-26 weeks, LWW = ℓn 20,
if 27-39 weeks, LWW = ℓn 33,
if 40-48 weeks, LWW = ℓn 44,
if 49-52 weeks, LWW = ℓn 51.

The logarithmic conversion is employed so as to be consistent with the dependent variable in the weekly wage rate earnings equations (LWR), which is equivalent to the difference between the logarithms of annual earnings and weeks worked (i.e., LWR = E - LWW). This specification allows testing of the proposition implicit in the estimation of the weekly-wage-rate model, that the coefficient of LWW is unity in the annual earnings model.

4.4. Market Effects Variables

4.4.1. Location

Where labor is less than entirely mobile geographically, shifts in labor from low real-wage areas to high real-wage areas will not eliminate inter-regional wage differentials for given skill levels. Real wage differentials induced by differences in the demand for labor relative to the supply within regional and local labor markets will tend to prevail for long periods of time. Even where perfect labor mobility would tend to equalize real-wage rates up to a cost of migration differential, differences in the amenities and costs of living associated with different regions and sizes of community will prevent the relative equalization of nominal wage rates.

Thus, geographic location is likely to have an independent effect on observed nominal earnings after all other factors are considered. Two sets of dummy variables are used to capture the effect of geographic location in the analysis, region and community type. Four regional dummies were constructed from the Geographic Code Field representing the commonly used regional divisions of the country - the Maritime Provinces (REG1), Quebec (REG2), Ontario (REG3), and the Prairie Provinces (REG4). British Columbia is used as a reference region. Community type is described by three dummy variables taken directly from the Place of Residence 1971 Field. These variables are: urban area of 30,000 persons or more (RES1); rural non-farm (RES2); and rural farm (RES3). The excluded category is urban areas of less than 30,000 persons.

4.4.2. Occupation

The type of work people do is broadly described by their occupation. Movement between occupations is limited to a greater or lesser degree by the differing skill and knowledge requirements of each occupation, as well as by a variety of institutional barriers such as licensing and certification requirements. Thus, both nationally and locally, each occupation or group of related occupations (as defined by required skills and knowledge) can be viewed as constituting a distinct labor submarket. Given the limitations to mobility between these sub-markets, systematic occupational differentials in nominal earnings can be expected to exist which reflect the differing demand situations and conditions of work characterizing occupational submarkets. A set of 16 occupational dummy variables defined directly(6) from the PUS Occupational Field was employed to capture these market effects nationally, and a continuous variable representing the interaction of province and occupation was employed to control for differences in occupational demand and conditions of work across the country.

The occupational dummies represent groupings of individual occupations that are reasonably homogeneous with respect to the type of work performed.(7) They are: Managerial, Administrative and Related Occupations (OC1); Natural Science, Engineering and Mathematics Occupations (OC2); Social Science and Related Occupations (OC3); Occupations in Religion (OC4); Teaching and Related Occupations (OC5); Medicine and Health Occupations (OC6); Artistic, Literary, Recreation and Related Occupations (OC7); Other Occupations (OC8);(8) Sales Occupations (OC9); Service Occupations (OC10); Occupations in Farming, Horticulture, and Animal Husbandry (OC11); Other Primary Occupations - fishing, hunting, trapping, forestry and mining (OC12); Processing Occupations (OC13); Occupations in Machining, Product Fabrication, Assembling, and Repairing (OC14); Construction Trades Occupations (OC15); and Transport Equipment Operating Occupations (OC16). The excluded category consists of Clerical Occupations.

The continuous variables used to capture inter-regional differences in labor market conditions facing particular occupations are the 1970 provincial mean male and female earnings prevailing for each individual's occupation in his or her

---

See footnote(s) on page 68.

province of residence (WAGE).[9] It should also be noted, that from the provincial perspective WAGE can be viewed a proxy for those supply-demand conditions which make the occupation in which an individual is working either a high or low paying occupation in the province of residence.

4.4.3 Industry

The demand for a particular type of labor is a derived demand that depends upon the characteristics of the production processes which employ that labor, the prices of the outputs of those production processes (hence, the demand-supply conditions for those outputs), and the quantities of the other inputs used in those production processes. Occupational designation should capture some of these demand-determining factors, region and community type should capture others. The bulk of the remaining factors are likely to be associated with the industry in which an individual is employed. Therefore, a set of dummy variables taken directly from the PUS Industry Field was added to the analysis. These variables are: Agriculture (IN1); Forestry (IN2); Fishing and Trapping (IN3); Mines, Milling, Quarries, and Oil Fields (IN4); Construction (IN5); Transportation, Communication, and other Utilities (IN6); Trade (IN7); Finance, Insurance and Real Estate (IN8); Community, Business and Personal Services (IN9); and Public Administration and Defence (IN10). Manufacturing serves as the excluded category.

This industrial taxonomy is rather coarse. It lumps together what may be very different sorts of productive activities, which vary from place to place in terms of the outputs they produce, and the input mixes they use to produce them. Therefore, the industry dummies were supplemented by an industry-province interaction in the form of a continuous variable measuring the 1970 capital-to-labor ratio of the industry of employment in the province of residence (KLR).[10] Besides being a convenient device for interacting province and industry, the effect of the capital-to-labor ratio on earnings has a straightforward economic interpretation. It will generally be the case that the more capital there is per worker in a particular employment setting, the higher will be the worker productivity and consequently the demand for labor services, with the result that if other things are equal, wages will be higher.

See footnote(s) on page 68.

Just as the capital intensity of the work setting is expected to affect worker productivity and earnings directly, it is also expected to affect the productivity of on-the-job training time (experience). In order to take account of such indirect effects on earnings KLR is combined with the experience variables in interaction terms.

4.5. Human Capital Characteristics

4.5.1. Formal Education

As discussed in the preceding chapter, individuals for the most part acquire those skills and knowledge that constitute their marketable human capital either in school or on the job. This study employs a number of variables to describe the school-acquired human capital an individual possesses. Dummy variables are used to describe broad ranges of educational attainment and whether this education had an academic or vocational focus. Continuous variables are used to measure actual years of school attended and quality of schooling.

The principal school attainment variable used in this study is YSCH. It is a continuous variable which measures years of schooling attended and is constructed from three PUS data fields - Level of Schooling, Length of Course or Apprenticeship, and Place of Highest Grade of Elementary or Secondary School. For highest levels of schooling attended below university a specific number of years of school attendance is assigned to each code in the Level of Schooling Field as follows:

No schooling	YSCH = 0.0 years
Below Grade 5	YSCH = 2.5 years
Grades 5-8	YSCH = 6.5 years
Grades 9-10	YSCH = 9.5 years
Grade 11	YSCH = 11.0 years
Grade 12	YSCH = 12.0 years
Grade 13	YSCH = 13.0 years

Since the number of years required for senior matriculation and for a bachelor's degree varies by province, a more complex procedure is used to assign year values to university levels of schooling attainment. Different YSCH values are assigned to a particular university level PUS code depending on the province in which the individual was assumed to have attended university, which was assumed to be the same province where he or she attended secondary school. The values assigned are given in Table 4.1.

While this procedure is an attempt at greater accuracy, it should be realized many of the assigned values are quite arbitrary. The sub-sample used in the analysis could contain individuals with the same university attainment level in cohorts as much as 40 years apart. Over the years changes have occurred in most provinces' educational systems with respect to years of school required for university admission and to earn a bachelor's degree. Provinces differ in their mix of three- and four-year degree holders and holders of advanced degrees. Quebec has two distinct school systems each implying different numbers of years of school at university levels of attainment.(11) Besides, university may have been attended in a province other than the one where secondary school was attended and advanced degrees obtained in yet a third province. Thus, any set of years of schooling assignments such as these, based on the codes in the PUS Level of Schooling and Place of Highest Grade of Elementary or Secondary School Fields, is likely to be in error for large numbers of individuals. The values used in this study represented the best collective wisdom available to the authors as to what set of values would be most representative of the years of schooling for the products of each province or territory's school system.

When it was indicated that an individual had completed a full-time vocational course other than an apprenticeship, the value of YSCH was increased by an amount depending on the duration of the course as indicated in the Length of Course or Apprenticeship Field. YSCH was increased by 0.3 years where the vocational course was 3-5 months long, 2.0 years where the course was six months to three years long and 3.5 years where the course was more than three years long.

The PUS Level of Schooling Field also serves as the basis for three variables describing general levels of attainment. The first of these variables (HS) indicates that the highest level of schooling attended was Grade 9 through Grade 13, that is, high school; the second variable (UN) indicates that the highest level attended was university. The excluded references category for these two variables is not having reached Grade 9. A third dummy variable (DEG) differentiates those

---

See footnote(s) on page 68.

TABLE 4.1. Values Assigned to YSCH

Location	Number of years of university education				
	1-2 years	3-4 years without degree	3-4 years with degree	5 or more without degree	5 or more with degree
Newfoundland	12.0	14.0	15.0	16.0	17.0
Prince Edward Island	13.0	15.0	16.0	17.0	18.0
Nova Scotia and New Brunswick	13.0	15.0	15.5	17.0	18.0
Quebec(1) and Ontario	14.0	16.0	16.5	18.0	19.0
Manitoba, Saskatchewan, Alberta and Northwest Teritories	13.0	15.0	15.5	17.0	18.0
British Columbia and the Yukon(2)	14.0	16.0	17.0	18.0	19.0
Outside Canada	13.0	15.0	16.0	17.0	18.0

(1) The values assigned individuals who completed secondary school in Quebec based on the assumption that students go from Grade 11 to a classical college for two years before enrolling in university.

(2) For British Columbia and the Yukon senior matriculation was assumed to be Grade 13 and that a bachelor's degree required four years at university.

Source: Based on discussions with University of Western Ontario Admissions Officers and personnel from the Education Liaison Branch of the Department of the Secretary of State and Statistics Canada.

who attended university, but did not receive a degree from those who did receive a degree. DEG is assigned a unit value where the individual has a university degree, zero otherwise. Having a university degree is expected to have a positive effect on earnings over and above having the equivalent years of schooling without a degree. This is popularly referred to as the "sheepskin effect". Although having a degree is expected to increase initial earnings levels, it is not clear whether this stems from the fact that the degree serves as a market signal for having attained specified levels of competence, or whether not having a degree signals certain negative personality attributes.

The above variables refer primarily to academically-oriented schooling. Two dummy variables based on the PUS Completed Full-Time Vocational Course and Length of Course or Apprenticeship Fields are used to describe participation in two different types of strictly vocational training programs which are not registered in the Level of Schooling Field. The first (COL), is coded unity to indicate completion of a vocational training course, other than an apprenticeship course, of three months or longer duration. It is hoped that this variable would capture the earnings effects of having attended a program such as those offered by the Colleges of Applied Arts and Technology. The second variable (V) indicates another sort of training program: it is coded unity where the individual reported having completed an apprenticeship course lasting six or more months. Apprenticeship courses are assumed to be less training-intensive per unit of time, and hence, require a longer duration than other types of vocational programs to have an equivalent effect on earnings. There is no adjustment in the YSCH variable for apprenticeship courses.

Equal years of schooling will result in differing amounts of accumulated human capital depending on the inputs the student brings to the education process and on the inputs the school provides. In part, at least, the student-supplied inputs should be captured by some of the personal attribute variables already described. A continuous schooling quality variable (SQ) was used to control for differences in school-supplied inputs. This variable measures in constant 1971 dollars the operating-expenditures-per-student-enrolled in the province and measures for the vintage year in which the individual attended the highest level of elementary or secondary school. The province was identified from the PUS Place of Highest Grade of Elementary or Secondary School Field. The vintage year was computed from the PUS Age Field and years of schooling variable YSCH discussed above.(12) The annual series of operating-expenditure-per-student-enrolled for each province were constructed by piecing together historical provincial current operating expenditure and enrollment data.(13) The annual cost data were converted to 1971 constant dollars.(14)

The school quality variable was assigned a value of zero for all individuals who did not attend school. However, because of data unavailability, it was also assigned a zero value for individuals whose Place of Highest Elementary or Secondary School was the Northwest Territories, the Yukon, or Outside Canada. To differentiate between non-school attendance and lack of data, a dummy variable (SQD) was used. This variable was coded one for individuals who attended school but for whom school quality data were not available.

4.5.2. Experience

Marketable human capital acquired on the job is expected to increase with the amount of work experience an individual has attained. The census questionnaire did not inquire about an individual's actual work experience. However, a potential experience variable can be constructed which should serve as a reasonably good proxy for actual experience where work activity is fairly continuous after the termination of formal schooling. This potential experience variable (EX) is computed as the individual's age less the number of years of schooling attended (YSCH), less 6.(15) The square of this variable ($EE=EX^2$) is also used in the analysis. It is employed as a device for allowing on-the-job accumulation of human capital, and hence earnings, to increase at a non-constant rate with additional experience (Mincer, 1974).

While this potential experience variable is a reasonable proxy for actual work experience where work activity is fairly continuous, as is generally the case with males, it is less appropriate for females whose work activity is frequently interrupted by the demands of family life - namely bearing and rearing children. In this regard the birth of a child signals a hiatus in a female's work activity. The length of this hiatus will vary with how early preceding pregnancy she leaves

---

See footnote(s) on page 68.

work, and how late in the process of child rearing, if ever, she returns to work. These absences from work can be expected to have a significant effect on a woman's earnings. During these periods she is not acquiring additional human capital through on-the-job training, and furthermore, the human capital she has already acquired is depreciating from lack of use. Therefore, two continuous variables derived from the PUS Number of Children Ever Born Field are used to augment the potential experience variables in the analysis of female earnings. The first of these is simply the number of children ever born (NCB), and the second is the number of children ever born squared (NC2). The squared term is introduced to allow the effect of an additional child on work experience, and hence on earnings, to vary with the number of children the woman has already had. These variables are used to make adjustments in the potential work experience variables using a two-stage procedure outlined in Chapter 3 and Appendix C.

FOOTNOTES

(1) The estimates of the male earnings equation presented in Chapter 5 were based on a 1 in 6 subsample of the males satisfying the selection criteria. This sampling ratio was selected because it would produce a subsample of approximately 10,000 observations, which was the largest sample size that could be conveniently handled by the procedure used to estimate the male earnings equations. The graphs in Chapter 2 are based on the entire sample.

(2) For those unemployed during that week, these characteristics describe the job of longest duration in 1970. Several different jobs may have been held during that year which affect earnings but which have not been documented

(3) The reason is that it is not possible to obtain a unique set of parameter estimates unless at least one category is excluded from each set of categories.

(4) Of course, a working spouse could be an otherwise unavailable source of financial assistance tending to increase educational attainment by reducing the opportunity cost of funds.

(5) This field differentiates between full- and part-time weeks only, not allowing the construction of a corresponding over-time variable.

(6) Individuals who had not stated their occupation were excluded from the subsamples used in the analysis.

(7) For an explanation of the occupational classification system used in constructing these groupings see Statistics Canada (Cat. No. CS12-536).

(8) This category is composed of materials handling and related occupations, other crafts and equipment operating occupations (largely in printing and utilities) and occupations not elsewhere classified.

(9) Explicit interaction of each occupational category with each province would produce an unmanageable number of variables. One hundred and forty-four additional variables would be required. Values for the WAGE variable were taken from Statistics Canada (Catalogue No. 94-769, Vol. III - Part 6).

(10) The capital stock figures were constructed from unpublished 1970 mid-year net fixed capital stock data for each industrial category by province, in current dollars, obtained from the Capital Stock Section, Construction Division of Statistics Canada. Employment data are from Statistics Canada (Catalogue No. 94-747, Vol. 3 - Part 4).

(11) The French educational system was used to assign values to YSCH for students educated in Quebec.

(12) The following formula was used to approximate the vintage year: Vintage = (1970 - AGE + 6 + YSCH/2), rounded off to the nearest whole year.

(13) The source of these data was the Education Division of Statistics Canada. For certain provinces in certain years these data were not reported; in these instances interpolation was used to fill in the gaps.

(14) Historical costs were inflated to constant 1971 dollars using Statistics Canada's Implicit Price Deflator for Expenditures on Consumer Goods and Services. The Implicit Price Deflator was extended back beyond 1926 by splicing the Wholesale Price Index for all Consumer Goods onto it.

(15) If this difference is negative, EX is set equal to zero. Conceivably this method of constructing potential experience could produce absurd results for the very young with little schooling. But given provincial mandatory school attendance requirements and the point system used for immigrant entry, instances of this are very infrequent in the sample.

CHAPTER 5

THE EMPIRICAL ANALYSIS OF MALE EARNINGS

5.1. Introduction

Three differently specified earnings equations were estimated for males. Two of these use the natural logarithm of annual earnings (E) as the dependent variable, the other uses the natural logarithm of the weekly-wage rate (LWR). One of the annual earnings equations, the full model, employs the complete set of explanatory variables--personal attributes (including the natural logarithm of weeks worked, LWW), human capital characteristics, and market effects. The other annual earnings equation, the reduced model, includes all of these same variables except for the market effects variables. The weekly-wage rate equation contains the same explanatory variables as the reduced annual earnings model, but excludes the logarithm of weeks worked which is implicitly part of the dependent variable (LWR = E - LWW). This model is subsequently referred to as the wage-rate model.

All three models were estimated with the maximum likelihood procedure outlined in Appendix B. Coefficient estimates for all three models are presented in Appendix E. The discussion in this chapter is restricted to a discussion of the principal implications of the coefficients. The coefficients of the full model are discussed in detail in Appendix E.

A statistical test of the hypothesis that the reduced model is the correct model was rejected at the 0.1% significance level. This suggests that market effects variables do, in fact, contribute significantly to the explanation of annual earnings over and above what is explained by the personal attribute and human capital variables. On the other hand, a test of the hypothesis that human capital and personal attribute variables contribute nothing to the explanation of individual earnings after market effects have been included in the model was also rejected at the 0.1% significance level. In other words, the human capital and personal attribute variables contribute significantly to the explanation of individual earnings even when such variables as industry, occupation, region, and average occupational wage rate, have already been included in the model.

It was indicated in Chapter 3 that if the estimated coefficient of the logarithm of weeks worked (LWW) is restricted to equal one in the annual earnings model, this model would be exactly equivalent to the wage rate model. However, both annual earnings models were estimated without restricting the coefficients of LWW, and the resulting estimated values were 0.85 and 0.83 for the reduced and full models respectively. Since both of these coefficients were significantly less than unity, this implies that annual earnings increase less than proportionately with weeks of work. In other words, those individuals who work less than a full year tend to receive a higher weekly wage rate than those who work a full year. This possibly reflects a weekly wage rate premium for workers in seasonal employment, or possibly a backward-bending labor supply phenomenon.

However, in spite of the fact that the estimated coefficient of LWW was significantly less than one in the reduced annual earnings model, the corresponding coefficient estimates for all other variables in the reduced annual earnings and weekly wage rate models were virtually identical.(1) This implies that factors such as ethnic group, level of schooling, etc., affect weekly earnings and annual earnings in a very similar way. Thus, the discussion in the remainder of this chapter is largely restricted to the full and reduced annual earnings models.

The coefficients of the regional variables are all significant and their magnitudes indicate that, <u>ceteris paribus</u>, earnings are highest in British Columbia, followed by Ontario, the Prairie Provinces, Quebec, and the Maritime Provinces, in that order. These significant coefficients support the proposition that after one has controlled for interpersonal differences in labor productivity as well as differences in industrial and occupational demand, labor supply does not adjust so as to equalize nominal earnings. These differences may reflect regional cost-of-living differences, interregional earnings differences that are not sufficient to warrant migration, or interregional labor immobility.

The coefficients for the community type variables indicate a significant relationship exists between community type and earnings. After controlling for all other effects, earnings were highest for those individuals living in urban areas

---

See footnote(s) on page 88.

with 30,000 or more population, followed in decreasing order by individuals living in urban areas with population less than 30,000 people, rural farm communities, and rural non-farm communities.

None of the migration variables turned out to have significant coefficients, which implies that after one has controlled for factors such as ethnic group, occupation, industry, schooling, experience and hours of work, recent internal migration and period of immigration from other countries have no significant residual effect on earnings.

The remainder of this chapter is devoted to an examination of the factors affecting earnings levels observed among males belonging to the various ethnic groups in the Canadian population, and to a general analysis of the components of earnings inequality observed among all males.

5.2. Ethnic Group Earnings Levels

The age-earnings profiles presented in Chapter 2 (Chart 2.2) show marked differences in the level (geometric mean) of earnings by age received by members of the various ethnic groups. In terms of the models used in this analysis, these differences can arise from differences in the levels of and/or rates-of-return to human capital as measured by years of schooling and experience, or from other factors not related to human capital. Table 5.1 indicates the average years of schooling and experience of males in each ethnic group as well as their geometric mean earnings levels. The Jews have by far the highest mean earnings. They also have the highest level of education and the second highest level of experience. At the other end of the spectrum, Native Indians have by far the lowest mean earnings and the lowest level of education. This group ranked fifth in terms of experience.(2)

The Negro/West Indian group has by far the lowest level of experience and the second lowest earnings level. They have, on the other hand, the second highest level of education. This paradoxical finding as well as what appear to be

---

See footnote(s) on page 88.

TABLE 5.1. Levels of and Marginal Rates of Return to Schooling and Experience and Earnings by Ethnic Group

Ethnic group	Percent of sample	Mean years of schooling	Mean years of experience	Geometric mean earnings	Full model		Reduced model	
					Rate of return to schooling	Rate of return to experience(1)	Rate of return to schooling	Rate of return to experience(1)
French	26.7	9.41	21.67	4,447	4.27	0.44	5.16	0.98
Negro/West Indian	0.3	11.64	15.27	4,105	(2)	(2)	(2)	(2)
North European	11.4	10.41	21.92	4,470	4.44	0.45	5.64	0.98
East European	6.1	10.12	24.06	4,629	4.83	0.28	6.20	0.78
Italian	3.9	8.24	23.01	5,014	4.40	0.12	4.62	0.37
Jewish	1.6	12.21	23.69	5,597	5.99	1.73	6.61	2.32
Oriental	0.8	11.35	20.13	4,402	4.82	1.09	7.28	1.75
Native Indian	0.7	7.31	21.72	2,465	1.07	-1.03	1.38	- .05
Other, unknown	4.1	10.15	20.66	4,817	4.03	1.11	4.65	1.50
British	44.6	10.85	22.11	4,770	5.82	0.72	7.47	1.31
Weighted average (3)		10.24	22.08	4,749	5.02	0.60	6.28	1.14

(1) Evaluated at the average experience level for the ethnic group.

(2) The calculation for the Negro/West Indian group are based on only 27 observations. Since some of the results seem to be unreasonable, this group has been dropped from subsequent discussion and analysis.

(3) Excludes the Negro/West Indian group.

Source: Computed from Tables E.2 and E.3 and Public Use Sample Tape.

unreasonable schooling and experience coefficients estimated for this group are probably attributable to the fact that the estimated for this group are probably attributable to the fact that the Negro/West Indian group contained only 27 observations. Because of this, and a particular concern about measurement error relating to the Negro and West Indian categories, this group is excluded from all subsequent interpretative analysis and discussion.

The simple correlation between average education level and geometric mean earnings is 0.59 across ethnic groups, and between average experience level and geometric mean earnings it is 0.36.

The rate-of-return to schooling and experience computed for each ethnic group from the coefficients estimated for the full and reduced models are also presented in Table 5.1.(3) The rate-of-return-to-schooling figures for the various groups indicate the percentage increase in earnings that an average individual in the group could expect from having one additional year of schooling. For example, for males of French descent, the full model estimate indicates that there would be an increase of 4.27% in earnings for each additional year of schooling, ceteris paribus. The comparable estimate for the reduced model is 5.16%.

The rates-of-return to schooling reported in Table 5.1 are for individuals who did not participate in an apprenticeship program and who did not marry prior to completing school. For all groups, the figures for individuals who married prior to completing school should be adjusted upwards by adding 3.50 to the full model and 3.74 to the reduced model. Similar adjustments should be made for all individuals who participated in apprenticeship training programs by subtracting 1.24 and 1.40 from the full and reduced model figures respectively. The positive adjustment for marriage prior to school completion is consistent with the hypothesis suggested in Chapter 4 that, given the increased opportunity cost of funds for married males, schooling will be undertaken after marriage only if it has an "above average" rate-of-return. The lower rates-of-return to schooling for individuals with apprenticeship training indicate that the specific training involved in apprenticeship program tends to reduce the economic value of formal education.

The figures for rate-of-return to experience presented in Table 5.1 have an interpretation similar to those for schooling.(4) Unlike the schooling figures,

---

See footnote(s) on page 88.

however, the rates-of-return to experience are not constant but decline with each additional year of experience. The figures reported in Table 5.1 have been computed using each group's average experience levels. Thus, for example, a typical male in the French group with the group's average level of experience can expect to get an income increase of 0.44% for an additional year of experience based on the full model estimate. The comparable estimate for the reduced model is 0.98%. The return-to-experience figures reported in Table 5.1 have been computed for individuals who did not marry prior to completing school, were not self-employed, did not participate in an apprenticeship program, and did not attend high school or college. The effects of these factors on the experience/earnings profile and, consequently, on the rate-of-return to experience depends on the level of experience. However, individuals who attended school after marriage, had apprenticeship training, or attended high school or university tend to have higher percentage increase in earnings per additional year of experience during early years of work, but these percentage increases tend to decrease relatively more quickly with labor force experience. Individuals who obtained a university degree have higher percentage rates of increase in earnings per year of experience than non-degree-holders throughout their working life. Self-employed individuals, on the other hand, will get lower percentage increases in earnings in early years relative to individuals working for wages or salaries.(5)

The rates-of-return to schooling as measured by the reduced model are higher than those measured by the full model for all nine of the ethnic groups. The weighted average rate-of-return to schooling in the full model is 5.02%, whereas in the reduced model it is 6.28%, the latter being 25.1% higher than the former. The rate-of-return to experience for each ethnic group computed from the reduced model is also higher than that computed from the full model. The weighted average rate-of-return evaluated at the overall average experience level implies a 0.60% increase in earnings per year of experience in the full model, and 1.14% increase in earnings per year of experience in the reduced model, the latter being 90.0% higher than in the full model. Attaining more education or experience makes it possible to obtain jobs in higher paying occupations and industries. Thus, the inclusion of occupational and industrial control variables lowers the observed rates-of-return to schooling and experience since these attributes of human capital

---

See footnote(s) on page 88.

are no longer picking up the job shifts accompanying education and experience increases.

Comparing ethnic groups in Table 5.1 it is apparent that groups with high marginal rates-of-return to schooling also tend to have high average levels of schooling. The marginal rates-of-return to schooling by ethnic group are plotted against average level of schooling by ethnic group for the reduced model in Chart 5.1. The simple correlation between the two is .83 for the reduced model and .87 for the full model. It appears that individuals in groups with higher rates-of-return to schooling respond by acquiring higher levels of schooling, and vice versa.(6) The net effect of this is that the intergroup variation in the level of schooling attended and its rate-of-return reinforce each other in contributing to the earnings differentials observed among the ethnic groups.

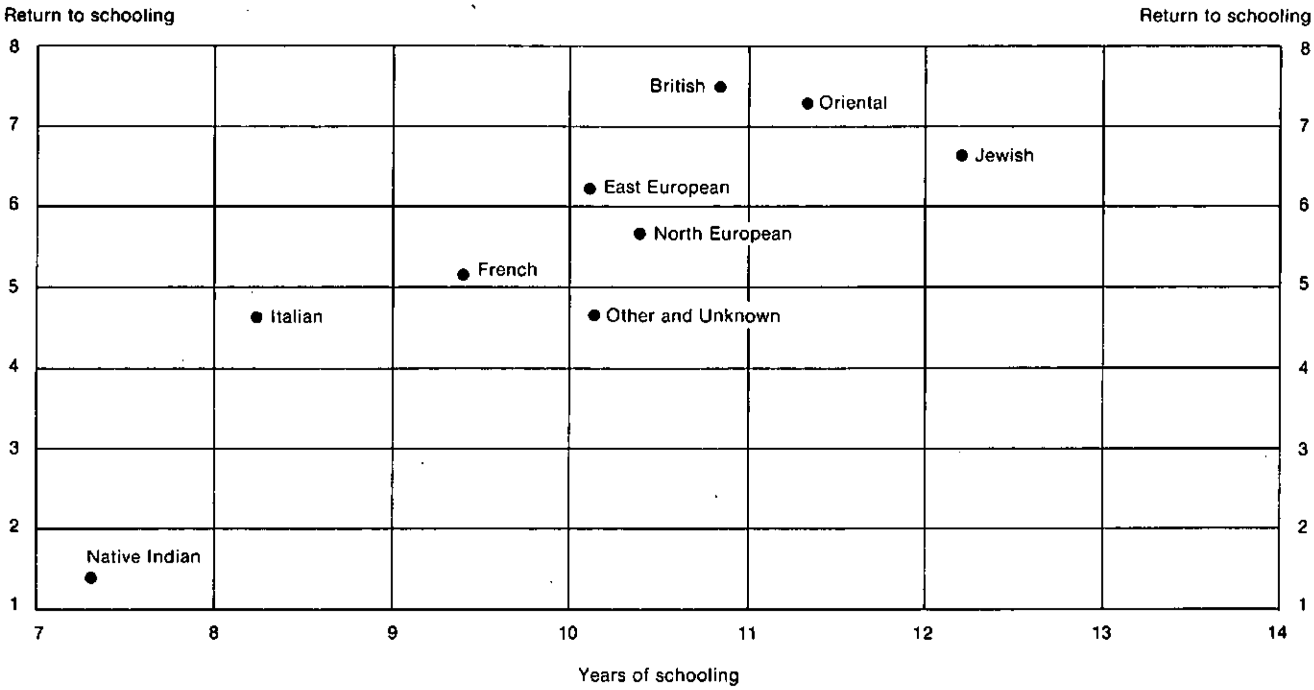
What is not obvious, however, is why the rates-of-return differ so markedly among the groups, varying from 1.38% for Native Indians to 7.47% for the British (based on the reduced model). These rate-of-return differences could arise from any one or more of the following differences among the ethnic groups: differentials in motivation for, or skills at, learning; differences in the quality of school attended;(7) and differentials in motivation or skills relating to the use made of acquired human capital in earning income. All of these differences are associated with labor supplied. It is also possible that differential rates-of-return are due to demand-side phenomena, namely racial and ethnic discrimination practiced by employers.

From the policy standpoint, it is interesting to compare the relative importance of differences in levels of schooling and experience, differences in rates-of-return, and residual ethnic effects in explaining earnings differentials among the various groups. Table 5.2 shows the effects on the ethnic groups' earnings levels of changing the levels of schooling and experience to the across-group average level while holding constant the rates-of-return at their actual levels; of changing rates-of-return to the across-group average rate-of-

---

See footnote(s) on page 88.

Chart — 5.1
Reduced Model Rates-of-return to Schooling Versus Years of Schooling by Ethnic Group



Source: Table 5.1

TABLE 5.2. Percentage Differences in Earnings by Ethnic Group for Alternative Combinations of Rates of Return and Levels of Schooling and Experience[1]

Ethnic group	Reduced Model					Full Model				
	Ethnic dummy (1)	Own rates, average levels (2)	Average rates, own levels (3)	Average rates and levels (4)	Total of average and dummy (5)=(1)+(4)	Ethnic dummy (6)	Own rates, average levels (7)	Average rates, own levels (8)	Average rates and levels (9)	Total of average and dummy (10)=(6)+(9)
French	20.4	- 4.7	-23.5	-29.3	- 8.8	16.8	- 3.7	-18.7	-23.2	- 6.4
North European	9.7	8.9	- 9.5	- 8.7	1.1	11.9	0.7	-10.1	- 9.4	2.5
East European	- 7.3	1.1	9.7	11.0	3.7	- 6.1	- 0.2	10.0	10.5	4.4
Italian	41.1	- 8.9	-31.4	-43.0	- 2.0	21.3	- 8.7	- 6.3	-25.9	- 4.6
Jewish	-34.3	17.1	72.6	86.8	52.5	-43.0	14.8	70.8	81.6	38.6
Oriental	-13.0	- 4.9	20.8	25.4	12.5	- 1.2	- 3.4	9.3	13.6	12.4
Native Indian	53.2	- 4.0	-56.0	-74.8	-21.7	54.6	- 2.8	-63.9	-78.9	-24.4
Other, unknown	7.5	- 2.4	-11.8	-14.1	- 6.5	- 4.5	- 1.8	- 1.7	- 3.3	- 7.7
British	-16.1	4.6	17.7	21.6	5.5	-11.6	3.6	12.4	15.5	3.8

(1) These percentages were derived as $100(b_j - \Sigma W_i b_i)$, where the b_i are the estimated regression coefficients of the ethnic dummies and the W_i are the proportion of the sample consisting of the ith ethnic group. These numbers should be interpreted as the percentage deviation of the geometric mean income of the jth ethnic group from the overall geometric mean that is not explained by the rest of the model.

Source: Computed from Tables E.2 and E.3 and Public Use Sample Tape.

return while holding constant levels of schooling and experience at their actual levels; and of changing both the levels and the rates-of-return to the across-group average level and rate. Finally, the table shows the percentage deviation of each group's mean earnings from the across-group mean of earnings that is explained by the ethnic dummy variable after controlling for human capital and all other factors.

The interpretation of the entries in Table 5.2 is as follows for the reduced model. Members of the French group tend to have a geometric mean income which is 20.4% above the overall average after controlling for all other effects. An individual with French rates-of-return to schooling and experience and the French levels of schooling and experience would receive 4.7% less than an individual with French rates-of-return and across-group average levels of schooling and experience. In other words, the "below average" levels of schooling and/or experience of the French group "account" for this group having an earnings level 4.7% below what it otherwise would be. On the other hand, an individual with French rates-of-return and French levels of schooling and experience would receive 23.5% less than an individual with across-group average rates-of-return and French levels. Thus, the "below average" rates-of-return of the French group result in 23.5% lower earnings. An individual who had French levels and rates would receive 29.3% less than an individual with both across-group average levels of and rates-of-return to schooling and experience. Finally, the "typical French individual" defined by the French dummy variable as well as French levels and rates-of-return is predicted to receive 8.8% less than the overall average as a consequence of "being French", all other things held constant. This percentage can be interpreted as the relative net effect of French ethnicity on earnings. The interpretations of the entries in the right half of the table for the full model are identical.

Comparing columns (2) with (3) and (7) with (8), it is apparent from both the full and reduced models that differential rates-of-return are far more important in explaining inter-ethnic group differences in earned income than are differences in levels of schooling and experience. Thus, it appears that the intergroup differences in skill production and/or imperfections in markets for the skills produced by schooling and experience result in intergroup differences in rates-of-return that are more important in determining ethnic earnings differentials than are differences in the distribution of levels of schooling and experience across ethnic groups. From a policy making standpoint this is unfortunate. Consider policies for raising the income levels of Native Indians. Based on calculations such as those in Table 5.2, raising the level of Indian education from its average of 7.31 years to the across-group average of 10.24 years would result in an income increase of only 3.1 or 4.0%, depending on whether full or reduced model estimates are used. On the other hand, if one could raise the rate-of-return for Indians' schooling and experience to the across-group average rates, it would result in earnings increases of 64 or 56%, depending on whether full or reduced model estimates are used. Raising the rates-of-return Native Indians receive, however, would be a much more difficult (if not impossible) objective to achieve than raising their schooling levels.

The combined effects of differential levels of and rates-of-return to human capital (columns 4 and 9) are at least partly offset for every group by other factors associated with the ethnic group that have not been explicitly controlled for in the models, and were consequently captured by the ethnic dummies. In every case, the effect of the ethnic group dummy (columns 1 and 6) was in the direction opposite to that of the total human capital effect (columns 4 and 9). Those groups having high levels of and rates-of-return to schooling and experience, such as Jewish and British (Table 5.1), tend to have large negative coefficients on their dummy variables (Table 5.2), while groups such as Italian and Native Indian which have low levels of and rates-of-return to human capital tend to have large positive coefficients on their dummy variables. Clearly, human capital plays a more important role in the determination of earnings for the former groups than it does for the latter. This conclusion holds for both the full and reduced model.

The inclusion of the market effects variables in the full model not only reduces the rates-of-return to schooling and experience, but it also tends to reduce (for six of the nine groups) the absolute magnitudes of the ethnic dummies. Thus, it appears that the presence of industrial, occupational and locational variables diminishes the residual effects of ethnicity and race per se on earnings. This is consistent with the hypothesis that certain ethnic groups have preferences for specific industries, occupations, or locations. But it is also consistent with the hypothesis that there is systematic discrimination practised by employers in certain occupations and industries that takes the form of direct barriers to entering higher paying industries and occupations, and these barriers are maintained regardless of the individual's qualifications in terms of education and experience.

5.3. Inequality Analysis

This section analyzes the sources of variation in the logarithm of the weekly-wage rate (LWR) and the logarithm of annual earnings (E). The logarithmic wage rate an individual receives is considered to be the sum of two types of effects--personal attributes and market effects (PME), and human capital effects (HCE). Thus, for the i^{th} individual

$$LWR_i = PME_i + HCE_i \tag{5.1}$$

Annual earnings are the product of the weekly-wage rate and the weeks of work. Hence, the logarithm of annual earnings is the sum of the logarithm of the wage rate plus the logarithm of weeks of work, a labor supply effect (LSE).(8) Thus, $E_i = LWR_i + LSE$. Substituting Equation 5.1 for LWR results in

$$E_i = PME_i + HCE_i + LSE_i \tag{5.2}$$

On the basis of Equation 5.1, the variance of the logarithm of the wage rate can be denoted as:

$$V(LWR) = V(PME) + V(HCE) + 2C(PME,HCE), \tag{5.3}$$

where V and C denote the variances and covariances of the bracketed terms.

The corresponding equation for annual earnings is

$$V(E) = V(PME) + V(HCE) + V(LSE) + 2C(PME,HCE) + 2C(PME,LSE) + 2C(HCE,LSE). \tag{5.4}$$

See footnote(s) on page 88.

Equation 5.3 and 5.4 can be used to partition the variances of the logarithm of weekly and annual earnings, which are the inequality measures being used in this study, into various sources. Thus, for example, inequality in the distribution of the wage rate among individuals arises from three sources: personal attribute and market effects, human capital effects, and an interaction of these two effects. Similarly, the inequality of annual earnings is partitioned into three direct sources: human capital effects, personal attributes and market effects, and labor supply effects. In addition, the interaction among these direct effects must be taken into account.

The variances of the logarithms of weekly and annual earnings have been partitioned into these sources using the method described in Appendix F. It should be noted that, as a consequence of the procedure used to estimate the earnings equations it is possible to allocate all of the variance to these sources, including that portion of the variance which was unexplained. The resulting variance partitions are presented in Table 5.3.

The variance partitions for all three models indicate that personal attribute effects are a more important source of inequality than human capital effects. (In the full model market effects are combined with personal attribute effects.) In every case, the direct personal attribute affects account for at least twice as much inequality as the human capital effects. Thus, the variations in the levels of and rates-of-return to human capital among individuals do not result in as much inequality as do variations in personal attributes (ethnicity, period of immigration, marital status, self-employment status, etc.) in the wage-rate and reduced models, and variations in personal attributes as well as market effects (occupation, industry, region, community size, etc.) in the full model. There is, however, a large negative covariance between these two sources of income inequality in all three models. This suggests that individuals who do relatively well in terms of earnings via human capital effects do relatively badly in terms of earnings via personal attributes and market effects, and vice versa. This implies that personal attribute and market effects and human capital effects tend to offset each other as sources of earnings inequality. In other words, if it were not for the fact that the interaction of these two sources tends to offset the direct effects, earnings inequality would be larger.

TABLE 5.3. Partition of the Variance of the Logarithms of Weekly-wage Rate and Annual Earnings(1)

	Personal attributes (and market effects)(2)	Labor supply effects(3)	Human capital effects
A. Wage-rate model: V(LWR) = .636			
Personal attributes	0.683		
Human capital effects	-0.252		0.205
B. Reduced annual earnings model: V(E) = 1.087			
Personal attributes	0.690		
Labor supply effects	0.196	0.217	
Human capital effects	-0.348	0.010	0.345
C. Full annual earnings model: V(E) = 1.087			
Personal attributes and market effects	0.655		
Labor supply effect	0.199	0.207	
Human capital effects	-0.249	0.011	0.284

(1) The elements on the principal diagonal are the variances and the off-diagonal elements are two times the covariance of the relevant effects. The components may not sum to the total due to rounding error.

(2) Market effects appear only in the full model.

(3) These effects are not present in the wage rate model.

Source: Computed from Tables E.1-E.3 and Public Use Sample Tape.

In both the reduced and full annual earnings models, the direct contribution of labor supply variations to earnings inequality is less than the direct contribution of either the personal attributes and market effects or the human capital effects. The covariance between labor supply effects and human capital effects is quite small and this interaction accounts for approximately 1% of the total inequality in earnings. However, there is a more substantial interaction between labor supply effects and personal attribute and market effects, and this interaction amounts to nearly 20% of the total inequality. Thus, it appears that individuals deriving high earnings from personal attributes, or personal attributes and market affects, tend to supply relatively larger amounts of labor. As a consequence, these latter two sources of inequality reinforce each other and tend to result in more inequality. On the other hand, the level of earnings from human capital does not seem to influence the amount of labor supplied, and these two effects do not reinforce or offset each other.

The partitions obtained from the reduced and full annual earnings models are quite similar. However, including the market effects controls in the full model results in a substantial reduction in the estimated contribution of the human capital affects as a source of inequality. The direct human capital effect contribution is 0.345 in the reduced model and 0.284 in the full model. There is an offsetting difference in the interaction between the human capital effects and the personal attributes and market effects. This interaction was -0.348 in the reduced model, whereas in the full model it was -0.294. Both of these changes can be attributed to the smaller estimates of the rates-of-return to human capital obtained in the full model. It is also interesting to note that the sum of the direct effects of personal attributes <u>plus</u> market effects in the full model is 0.655, which is <u>less</u> than the 0.690 direct contribution of only the personal attribute effects in the reduced model. Thus, the market effects tend to offset the personal attribute effects in the full model.

It is instructive to partition the variance of the logarithm of annual earnings another way. Substituting Equation 5.3 into Equation 5.4 yields

$$V(E) = V(LWR) + V(LSE) + 2C(LWR,LSE), \qquad (5.5)$$

where 2C(LWR,LSE) = 2C(PME,LSE) + 2C(HCE,LSE). Table 5.4 portrays the variance partitioning corresponding to Equation 5.5 for both the reduced and full annual earnings models. The results are almost identical for the two models. Approximately 62% of the variance in the logarithm of annual earnings is attributable to variations in the weekly-wage rate, 19% to the variations in labor supply, and 19% to the covariation between these two effects. Thus, one can conclude that inter-personal variations in weekly earnings are far more important than inter-personal variations in weeks worked in producing the observed inequality in individual annual earnings.

TABLE 5.4. Partition of the Variance of the Logarithm of Annual Earnings With Respect to Wage Rate and Labor Supply Effects[1]

	Wage rate effect	Labor supply effect
A. Reduced model: V(E) = 1.087		
Wage rate effect	.687	
Labor supply effect	.206	.217
B. Full model: V(E) = 1.087		
Wage rate effect	.690	
Labor supply effect	.210	.207

(1) The elements on the principal diagonal are the variances,and the off-diagonal elements are two times the covariances of the relevant effects. The components may not add to the total due to rounding error.

Source: Computed from Table 5.3.

FOOTNOTES

(1) Most coefficient estimates differ by less than one standard error between the two sets of estimates, and only one coefficient differs by more than two standard errors. This was the interaction of married-in-school (MIS) and experience (EX), and those coefficients differ by less than 2.3 standard errors.

(2) Given the high rate of unemployment among Indians, the potential experience measure used (age minus years of schooling minus six) probably overestimates actual work experience.

(3) The rates-of-return to schooling for the various groups are calculated as

$$100 \left(\frac{\partial E}{\partial YSCH}\right) = 100 \quad (a_i + b \quad \overline{SQ} \quad + c \quad \overline{KLR}),$$

where a_i is the estimate of the schooling coefficient for the ith ethnic group, b and c are the estimated coefficients for the interactions of years of schooling with school quality and with the capital/labor ratio, respectively, and SQ and KLR are sample means for the school quality and capital/labor ratio variables, respectively. (Note c is zero in the reduced mode since KLR is not included in that model.)

(4) The rates-of-return to experience for the various groups are calculated as

$$100 \quad \left(\frac{\partial E}{\partial EX_i}\right) = 100 \; (d_i \quad + \quad 2e_i\overline{EX}_i \quad + \quad f\overline{KLR} \quad + \quad 2g\overline{KLR} \quad . \quad \overline{EX}_i),$$

where d_i and e_i are the estimated coefficients for experience and experience squared for the ith group, f and g are the estimated coefficients for the interactions of KLR with experience and with experience squared, respectively, $\overline{KLR}$ is the mean capital/labor ratio, and $\overline{EX}_i$ is the mean experience level for group i. (Again, note that both f and g are zero in the reduced model since $\overline{KLR}$ is not included in that model.)

(5) The effects of these factors on the age-earnings profile are discussed in greater detail in Appendix E.

(6) A similar pattern is not evident with respect to experience and its rate-of-return. This is probably due to the fact that individuals have no control over their potential experience levels once they leave school. Hence, those individuals in groups with high rates-of-return to experience cannot deliberately acquire high levels of experience.

(7) Although we attempted to control for school quality by including an annual provincial expenditure per enrolled pupil variable (SQ), this would not control for variations within provinces associated with ethnicity or race of the student body. Besides, this variable's poor performance in the regressions indicate it is probably not a very good proxy for school quality in general.

(8) This labor supply effect incorporates the coefficient for LWW, which in the reduced model was 0.847 and in the full model was 0.827.

CHAPTER 6

THE EMPIRICAL ANALYSIS OF FEMALE EARNINGS

6.1. Introduction

This chapter discusses the factors affecting the earnings of individual females. It utilizes analyses of earnings functions estimated separately for five ethnic groupings of females.

One of the shortcomings of the PUS data is lack of information on labor force experience. For males, potential labor force experience can be calculated as age minus years of school attendance minus six, and this can be considered a reasonable estimate of actual experience. Females, however, tend to drop out of the labor force for childbearing and rearing. Consequently, the same procedure is less appropriate for women. In Chapter 3 and Appendix C, a two-stage estimation procedure was outlined which can be used to estimate hiatuses in work activity due to childbearing and rearing. This allows adjustments to be made in the potential experience variable for discontinuities in experience related to motherhood.

This two-stage procedure was used to estimate earnings functions for 31,481 individual females who satisfied the criteria specified in Chapter 4. In order to simplify computation and estimation, functions were estimated for each of five ethnic groupings, separately. These groups consisted of females claiming the following descent:

1. British Isles	15,093 individuals
2. French	7,623 individuals
3. North European (Austrian, Finnish, German, Netherlands, Scandinavian)	3,298 individuals
4. East and South European (Czeck, Hungarian, Polish, Russian, Slovak, Ukranian, Italian and Jewish)	3,592 individuals
5. Other Ethnic Groups (Negro, West Indian, Chinese, Japanese, Native Indian, Other, and Ethnic Group Unknown)	1,875 individuals

Three separate specifications were estimated for each earnings function. The results for the first stage for one of the specifications and the second-stage results for all three specifications are reported in Appendix G, Tables G.1 through G.4. The coefficient estimates of one of the specifications, the full annual earnings model, are discussed in detail in that Appendix. In the remainder of this section, we discuss the principal implications of the coefficient estimates. Interested readers are referred to Appendix G for a more detailed discussion.

The simplest specification, the reduced model has the logarithm of annual earnings as its dependent variable. This model includes continuous variables for years of schooling, adjusted potential labor force experience and the square of this variable, an estimate of the labor force hiatus, and the logarithm of weeks of work. In addition it includes dummy variables for having a university degree, working part-time hours, for ever having been married, and sets of dummies for region and for type of community.(1) Statistical tests indicated that each of the variables (or sets of region and community type variables) contributed significantly to the explanation of annual earnings. All variables had their expected signs with the exception of the ever-married variable. This variable indicated that women who were ever-married earned more income than never-married women in the same ethnic group after controlling for all other factors. This ever-married/never-married differential ranged from 10% for Other females to 26% for East and South European females. (Possible reasons for this surprising finding are discussed in detail in the next chapter.)

The variables for type of community indicated that, after controlling for all other factors, earnings of women were lowest in rural farm communities and were second lowest in rural nonfarm communities. Women in urban areas under 30,000 in population had the highest earnings with the second highest occurring in urban communities with populations of 30,000 and over. This ordering was the same for all ethnic groups. The variables for geographic region indicated no consistent pattern of regional variation in earnings among the ethnic groupings.

A second specification using the logarithm of annual earnings as the dependent variable was also estimated. This model is subsequently referred to as the

---

See footnote(s) on page 106.

full annual earnings model. The explanatory variables in this specification included all those in the reduced model. In addition, variables to control for the average wage in the occupation and province of employment, self-employment, being the head of a census family, working overtime hours, being proficient in both official languages, and the occupation of employment were included. Furthermore, the ever-married variable of the reduced model was disaggregated into two variables -- one for currently married and one for widowed, divorced, or separated. Significance tests indicated the additional variables contributed significantly to the explanation of annual earnings over and above the variables used in the reduced model.

The major surprise among these additional variables are negative and statistically significant coefficients for the variable for the French and North European groups. The overtime hours coefficient was positive and significant (as expected) for the British group, and it was not significantly different from zero for either the East and South European or the Other groups. These coefficients indicate that, after controlling for all other factors, females of French and North European descent who work overtime hours earn 6.9% and 11.4% less, respectively, than do women in these groups who work normal hours. Women in the British group who work overtime hours earn 4.7% more than their counterparts who work regular hours. The two variables for being currently or formerly married had positive coefficients for all ethnic groups indicating, as in the reduced model, that ever-married women tend to earn more than never-married women after controlling for all other factors. The ordering of earnings by type of community for each of the ethnic groups was the same as in the reduced model. As in the reduced model, the coefficients of the regional variables showed no consistent pattern among the ethnic groups; furthermore, these variables tended to be less significant in the full model than in the reduced model. This suggests that in the reduced model the regional variables serve as proxies for demand effects, whereas in the full model the occupation and the provincial-occupational wage variables capture some of these effects, tending to reduce the explanatory power of the regional variables.

The third specification estimated for each of the ethnic groups is designated the wage rate model. The dependent variable in this model is the logarithm of annual earnings divided by the weeks of work (i.e., the logarithm of the implied weekly wage rate). The explanatory variables in this model are identical with those of the full annual earnings model with the exception that the logarithm of weeks of work is not included.[(2)] The coefficient estimates for the wage rate model are virtually identical to those for the full model.[(3)]

6.2. Factors Affecting the Level of Female Earnings

The geometric mean earnings of females for the various ethnic groups were plotted against age in Chart 2.3. While there do not appear to be substantial differences in female age-earnings profiles by ethnic grouping, the importance of education in the determination of individual female earnings is clearly illustrated by the age-earnings plots by level of education in Chart 2.1. Those females with higher schooling levels tend to have higher incomes, especially those who are between 25 and 65 years of age.

The average years of schooling and geometric mean earnings by ethnic groupings are presented in Table 6.1. The North European group has the lowest, whereas the French group has the highest geometric mean earnings, the latter being 17% higher than the former. On the other hand, the North European females have the second highest average level of education (10.94 years) whereas the French females have the third highest level of education (10.25 years). The British females have the highest level of education (11.36 years), but the second lowest earnings level. At the other end of the education spectrum, the Other female grouping has the sec second highest earnings level but has the lowest level of schooling (10.02 years). The simple correlation between years of schooling and geometric mean earnings across these five groups is -0.76. We have no ready explanation for the negative correlation between geometric mean earnings and years of schooling across these ethnic groups. (The comparable figure for nine male groups was positive as expected and equalled 0.59.)

Within ethnic grouping, the partial effect of schooling level on earnings level was positive and significant. The magnitude of the effect is indicated in the last three columns of Table 6.1. These figures measure the percentage increase in earnings that an individual would experience from each additional year of schooling when all other factors are held constant.[(4)] For example, the estimate obtained from the reduced model for the British group implies that each additional year of

---

See footnote(s) on page 106.

TABLE 6.1. Geometric Mean Income, Years of Schooling, and Rates-of-Return to Schooling for Females by Ethnic Groups

Ethnic group	Geometric mean earnings	Years of schooling	Return to schooling		
			Full annual earnings	Wage rate	Reduced annual earnings
	$				
British	1,980	11.36	3.78	3.76	6.63
French	2,123	10.25	3.99	3.91	7.61
North European	1,821	10.94	3.68	3.68	6.18
East and South European	2,115	10.02	2.10	2.11	4.17
Other	2,122	10.02	2.36	2.33	5.10
Weighted average	2,019	10.81	3.54	3.51	6.45

Source: Tables G.2 through G.4 and Public Use Sample Tape.

schooling would, ceteris paribus, result in an earnings increase of 6.63%. The most remarkable feature of these figures is that the implied percentage increase in earnings per year of schooling in the reduced model is much higher than in the full model for all ethnic groupings, in some cases being nearly twice as large in the reduced model. Comparing the weighted average returns to schooling for all females from the full annual earnings model with the reduced annual earnings model, the return to education increases from 3.54 to 6.45, an 82% increase. This difference is probably due to the fact that higher paying occupations are only open to individuals with higher levels of schooling. Since the occupation of employment is held constant in the full model, earnings increases associated with higher education result only from higher pay within occupations. In the reduced model, however, since there are no controls for occupation, the return to education is a composite of the increase earnings associated with higher paying occupations and increased earnings within occupations.(5)

There was also considerable variation in the return to education across ethnic groupings, the French females having the highest and the East and South European females having the lowest return to education. The simple correlation between geometric mean earnings and return to education is also negative but much smaller in absolute magnitude than between years of schooling and earnings. On the other hand, the correlations between the rate-of-return to schooling and level of schoolings are positive for all three sets of rate-of-return estimates and equal to 0.7 for both of the full models and 0.5 for the reduced model. The rates-of-return to schooling for the full annual earnings model are plotted against the years of schooling in Chart 6.1. From this figure it is obvious there is a positive association between rate-of-return to and level of schooling, with the French group deviating somewhat from the pattern.

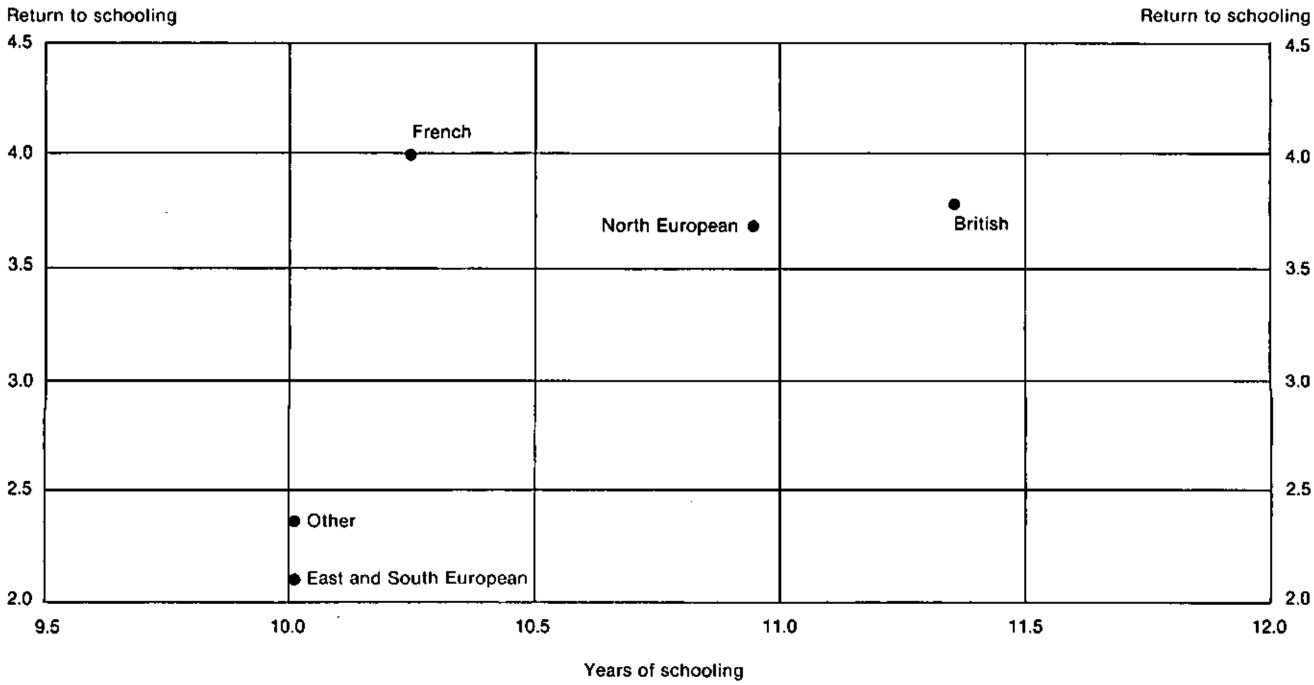
As relates to formal schooling being the cause of observed earnings differences among female ethnic groupings, we have the following situation: whereas rates-of-return to schooling are positively associated with educational attainment across ethnic groupings, both attainment and rates-of-return are negatively related to earnings levels across ethnic groupings. Thus it appears that ethnic differences in levels of schooling attained and the returns earned from this schooling are not

---

See footnote(s) on page 106.

Chart — 6.1

Years of Schooling Versus Full-model Rates-of-return to Schooling for Females by Ethnic Group



Source: Table 6.1

a major source of inter-ethnic earnings differentials for females. Other sources for these differentials must be explored.

Inter-ethnic geometric mean earnings differentials can be partitioned into two parts. The first is that part of the differential which can be explained by inter-ethnic differences in attribute levels (education, experience, etc.). The second part is a residual which can also be interpreted as the portion of the inter-ethnic earnings differential stemming from inter-ethnic differences in the rates-of-return to attributes.

The first part of the differential is calculated by taking the inter-ethnic differences in mean attribute levels and weighting these differences by the appropriate regression coefficients. Formally, it is calculated as $100 \sum_k (\bar{X}_k^1 - \bar{X}_k^2) b_k^i$, where $\bar{X}_k^i$ is the mean level of attribute k for ethnic groups i (i=1,2), b_k^i is the estimated regression coefficient, and the summation is over all the independent variables in the model. The resulting quantity is the percentage by which the earnings of Group 1 exceed those of Group 2 because of the inter-ethnic differences in attribute levels it can be negative). When comparing two ethnic groups, the magnitude of this term will depend on whether the regression coefficients of the first or second group are being used as weights (i.e., there is an index number problem). Consequently, both sets of regression coefficients should be used in every pairwise comparision as they will generate upper and lower bound estimates of the portion explained by attribute differences.

These upper and lower bounds can be subtracted from the percentage difference in earnings between the ethnic groups to obtain upper and lower bounds on the residual. This residual can also be calculated as the inter-ethnic differential in the regression coefficients weighted by the mean attribute levels of one of the ethnic groups as $100 \sum_k (b_k^1 - b_k^2) \bar{X}^i$. As above, the magnitude of the result depends on whether the weights (attribute levels) of Group 1 or 2 are used. Thus, upper and lower bounds on the rate-of-return effects are provided by using the attribute levels of the two groups as weights.[(6)] The partitioning procedure is discussed in greater detail in Appendix H.

See footnote(s) on page 106.

The procedure described above was used to partition the gross differences in geometric mean earnings between pairs of the five ethnic groupings. The results are reported in Table 6.2. Based on the signs of the attribute level effects for the various ethnic pairs, it is evident that the French female group had the highest weighted average level of attribute while the North European female group had the lowest weighted average level. This low level of attributes, in general, was the principal reason the North European females had the lowest earnings. Based on the signs of the rates-of-return effects, it is evident that the British females had the lowest weighted average rate-of-return on their attributes in general, and this is the reason the British females had the second lowest earnings level. The East and South European group, on the other hand, had the highest average rates-or-return, but the effect of these high rates-of-return on attributes was offset by the fact that this group had the second lowest level of attributes.

This section has focused primarily on differences in earnings levels among ethnic groupings. The next section discusses the degree of income inequality within ethnic groupings.

6.3. Inequality of Female Earnings

The measure of inequality used in this study is the standard deviation of the logarithm of earnings. As discussed in Chapter 3, the problem with this and all other measures of inequality is that they are meaningful only as relative measures. In other words, all that can be done is to compare the degree of inequality among various distributions to decide which are relatively more unequal. It will also be useful to reemphasize the fact that different inequaltiy measures can and do result in different renkings, and we are only examining one measure, the standard deviation of the logarithm of earnings.

The standard deviation of the logarithm of earnings for the five female ethnic groupings is reported in the first row of Table 6.3. These figures suggest that the North European and British groupings have the most unequal distribution of earnings, while the French, East and South European, and the Other groupings have somewhat more equal distributions. Thus, the North European and British groups not only have the lowest geometric mean earnings (Table 6.1), but these earnings are more unequally distributed than for the other three ethnic groupings. On the other

TABLE 6.2. Partitioning of Female Inter-ethnic Earnings Differences Based on Full Annual Earnings Regressions(1)

Groups compared	Gross differential(4)	Attribute level effects(2) Lower bound	Attribute level effects(2) Upper bound	Rates-of-return(3) Lower bound	Rates-of-return(3) Upper bound
British/French	- 7.0	- 3.7	- 4.9	- 2.0	- 3.3
British/North European	8.3	10.9	11.0	- 2.6	- 2.7
British/East and South European	- 6.6	1.3	2.2	- 8.0	- 8.9
British/Other	- 7.0	- 0.9	- 2.8	- 4.2	- 6.1
French/North European	15.3	10.5	16.9	- 1.6	4.8
French/East and South European	0.4	4.9	7.3	- 4.5	- 7.0
French/Other	0.1	3.2	3.3	- 3.2	- 3.3
North/East and South European	-14.9	- 7.7	- 9.9	- 5.0	- 7.2
North European/Other	-15.3	-11.9	-12.7	- 2.6	- 3.4
East and South European /Other	- 0.3	- 2.7	- 3.1	2.3	2.8

(1) The numbers in the table are percentages. Positive (negative) numbers indicate that the group listed first (second) in the row had the advantage.

(2) Computed as $100 \sum_k (\bar{X}_k^1 - \bar{X}_k^2) b_k^i$, where i(=1,2) indicates the first and second ethnic group in the row, $\bar{X}_k^i$ is the mean level of attribute k for ethnic grouping i; and b_k^i is the associated regression coefficient for ethnic group i.

(3) Computed as $100 \sum_k (b_k^1 - b_k^2) \bar{X}_k^1$, where all symbols are defined as above.

(4) Computed as $100(\bar{E}_1 - \bar{E}_2)$, where $\bar{E}_i$ is the mean of the logarithm of earnings for the ethnic group listed first (i=1) and second (i=2) in the row. The resulting number is the difference between the geometric mean earnings of the two groups as a percent of the average of the two means.

Source: Computed from Table G.2 and Public Use Sample Tape.

TABLE 6.3. Sources of Female Earnings Inequality Within Ethnic Groups

Source of inequality	Ethnic groups: British	French	North European	East and South European	Other
Total(1)	1.21	1.12	1.22	1.11	1.10
Years of schooling(2)	.18	.24	.18	.16	.21
University degree(2)	.05	.05	.07	.06	.05
Years of experience(3)(4)	.49	.54	.57	.36	.44
Weeks of work(2)	.69	.64	.70	.7	.63
Part-time work(2)	.36	.24	.34	.30	.26
Marital status(2)	.07	.07	.07	.11	.05
Labor force hiatus(2)	.05	.06	.09	.05	.05
Type of community(3)	.05	.06	.05	.05	.06
Region(3)	.06	.08	.06	.05	.05

(1) Computed as the standard deviation of the logarithm of annual earnings. The overall standard deviation for the five groups is 1.173.

(2) Computed as $|b_i|\ S_i$, where $|\ b_i\ |$ is the absolute value of the appropriate regression coefficient from Table G.4 and S_i is the standard deviation of the corresponding independent variable.

(3) Computed as $\left[\Sigma\ b_i^2\ S_i^2\right]^{\frac{1}{2}}$, where i ranges over the appropriate variables.

(4) These figures are overestimates. Lower bounds are as follows:
British .39; French .44; North European .46; East and South European .30; and Other .37. The true values lie somewhere between these values and those reported in the body of the table. For an explanation, see footnote 10 in the text.

Source: Computed from Table G.4 and Public Use Sample Tape.

hand, the French, East and South European, and Other groupings have relatively more equal distributions of earnings and also have higher earnings levels.(7)

An indication of the sources of income inequality for the various ethnic groupings can be obtained by multiplying the absolute values of the estimated regression coefficients from the earnings functions by the standard deviation of the corresponding independent variables. The resulting numbers indicate the contribution that variations in the explanatory variable make to variations in the logarithm of earnings (as measured by the standard deviation), all other factors held constant.(8) In the case where a set of variables is required to control for some factor (region, for example), only the total effect for the set is reported.(9) The results of these calculations are reported in Table 6.3.

The larger the numbers in a particular row in this table, the greater the contribution of that factor to inequality in the distribution of earnings. Based on this criterion, it is clear that the major source of earnings inequality for all of the ethnic groupings is the distribution of weeks of work. This is followed in importance by the inter-ethnic distributions of years of experience,(10) part-time work and years of schooling (in that order). Perhaps the most surprising feature of these results is the small magnitude of the contribution from the regional dummies. The quantities in this row are among the smallest in the table. This suggests that region does not contribute very much per se to income inequaltiy once all other factors are taken into account, and is consistent with our earlier discussion of the relative importance of the regional control variables in explaining levels of earnings.

In comparing the sources of earnings inequality across ethnic groupings, not much can be said about why some ethnic groupings have relatively more unequal earnings distributions that others. The reason for this is that many of the explanatory variables' effects exhibit considerably more variation across ethnic groups than do the inequalities in the earnings distributions themselves. This is particularly true of the inter-ethnic variations in the effects of schooling and marital status.

See footnote(s) on page 106.

To summarize, the three most important sources of intra-ethnic earnings inequality were found to be the distributions of weeks of work, part-time work, and years of potential labor force experience. Chapter 8 is devoted to an analysis of the distribution of weeks and hours of work. The distribution of years of potential labor force experience is discussed in the next section.

6.4. Estimated Years of Labor Force Experience

One of the major problems in explaining the earnings of women using the PUS data is that this data base does not contain information on their labor force experience. Females tend to drop out of the labor force for purposes of bearing and rearing children, and no data are available on the length of the hiatuses in work activity. We have attempted to estimate the length of these hiatuses on the basis of the number of children ever born using the procedure described in Chapter 3 and Appendix C. These estimates were used to adjust computed years of potential experience based on age and years of schooling. Recently, some data on the median years of labor force experience for males and females in different age groups have been published by Statistics Canada (Catalogue No. 13-557). The purpose of this section is to compare our estimates with those of Statistics Canada.

The most relevant Statistics Canada data are reported in Table 6.4. Comparable estimates using our procedures are reported by ethnic group in Table 6.5. One very important distinction between the data in Tables 6.4 and 6.5 must be noted: Statistics Canada has reported medians whereas we have reported means, and this makes comparisons between the two somewhat difficult.

Comparing our results for males with those reported by Statistics Canada, it is evident that our means for the 45 and over age groups are substantially higher than their median for the same age group. However, when comparing means and medians for open-ended classes of positively skewed distributions such as the ones considered here,(11) the median will always be lower than the mean. The medians and means for males in the lower age groups are much closer, which is consistent with a lesser degree of skewness in the lower age groups.

See footnote(s) on page 106.

TABLE 6.4. Median Number of Years in the Labor Force by Age and Sex, for Canada, 1972

Age	Median years in labor force		
	Male	Female	Difference
Less than 24	3.0	2.8	0.2
25-34	10.7	7.8	2.9
35-44	21.5	12.6	8.9
45 and over	33.8	18.2	15.6
Total	18.5	7.7	10.8

Source: Earnings and Work Histories of the 1972 Canadian Labor Force. Statistics Canada, Catalogue Number: 13-557.

Comparing the females results, the means and medians for the two lowest age groups are reasonable close. However, they deviate rather markedly for the upper two age groups. The female age distribution, however, exhibits a large amount of positive skewness.(12) This undoubtedly affects the results and perhaps makes the comparison for the upper age groups meaningless.

Nonetheless, it appears that our procedure may have substantially underestimated the length of hiatus from the labor force for older females. As a consequence of this, ever-married and never-married women aged thirty and over are analyzed separately in the following chapter. These results are then compared with results for ever- and never-married males thirty years of age and over.

See footnote(s) on page 106.

TABLE 6.5. Estimated Average Years of Labor Force Experience by Males and Females by Age and Ethnic Groups

Ethnic groups	Age group	Male	Female	Difference	Estimated female hiatus
British	Less than 24	2.8	2.0	0.8	0.6
	25-34	11.6	8.3	3.3	2.8
	35-44	22.9	17.8	5.1	4.7
	45 and over	39.7	34.1	5.6	4.0
	Average	22.2	16.4	5.8	2.8
French	Less than 24	3.9	3.2	0.7	0.3
	25-34	12.8	9.9	2.9	1.9
	35-44	24.4	20.4	4.0	3.7
	45 and over	40.6	35.8	4.8	3.4
	Average	21.8	14.8	7.0	1.9
North European	Less than 24	2.7	2.1	0.6	0.5
	25-34	12.1	8.5	3.6	2.8
	35-44	23.4	17.8	5.6	5.0
	45 and over	40.0	34.1	5.9	4.7
	Average	22.0	15.6	6.4	3.0
East and South European	Less than 24	2.9	2.1	0	0.6
	25-34	12.5	9.0	3.	3.3
	35-44	24.0	19.2	4.8	5.2
	45 and over	40.7	34.6	6.1	4.8
	Average	23.7	16.4	7.3	3.3
Other	Less than 24	4.0	2.8	1.2	0.6
	25-34	12.4	9.9	2.5	2.4
	35-44	23.7	19.4	4.3	4.2
	45 and over	40.2	35.2	5.0	3.9
	Average	20.4	14.6	6.8	2.5

Source: Computed from Public Use Sample Tape and the results of the first-stage regressions for the reduced annual earnings model. The method is described in Appendix C.

FOOTNOTES

(1) The exact variables included can be determined from Table G.4 of Appendix G. Variable names are defined in Appendix D.

(2) In essence, the two specifications are identical except that the coefficient of the logarithm of weeks of work is restricted to equal unity in the wage rate model.

(3) For the five ethnic groupings only six of 130 estimated coefficients differ by more than one standard error between the annual earnings and wage rate models, and none differ by as much as two standard errors.

(4) These figures are simply the appropriate regression coefficients from the regression multiplied by 100. That is, $100(\frac{\partial E}{\partial YSCH}) = 100b_{YSCH}$, where b_{YSCH} is the estimated coefficient for the years of schooling variable.

(5) A similar but less dramatic effect was also observed with the males. The comparable weighted average figures for the nine male ethnic groups were 5.02% for the full model and 6.28% for the reduced model, the latter being 25% higher (compared with an 82% increase for the females). This suggests that there is more occupational advancement associated with education for females than for males. However, within occupations, it appears that additional schooling has relatively greater financial rewards for men than for women.

(6) The regression coefficients are "rates-of-return" to the various attributes and some researchers have argued that differences in these rates-of-return between two groups arise from discrimination (Masters, 1974; Oaxaca, 1973; Robb, 1978). While some of these differences may be the result of discrimination, we feel that motivation and self-selection can also be major factors in explaining differences in coefficients between two population groups.

(7) For all pratical purposes, the geometric mean earnings of the French and Other groups are equal since they differ by only one dollar. Also, the degree of inequality of the French, East and South European, and Other distributions are virtually identical.

(8) This is calculated as $|b_i| S_i$, where $|b_i|$ is the absolute value of the estimated regression coefficient and S_i is the sample standard deviation of the independent variable. The result differs from a "standardized" regression coefficients which is defined as $b_i S_i / S_y$, where S_y is the standard deviation of the dependent variable. We have not standardized by S_y since we are interested in allowing the factors which contribute to the S_y to be larger for some ethnic groupings than others.

(9) This total effect is calculated as $\sum_i b_i^2 S_i^2 {}^{\frac{1}{2}}$ where i ranges over all the variables in the particular set. This corresponds to the standard deviation for the set. The covariances among the dummy variables within a particular set are all zero because of the way these variables are constructed.

(10) The figures for years of potential work experience reported in the body of the tables are overestimates. The correct calculation would be given by $\left[b_{EXS}^2 S_{EXS}^2 + b_{EES}^2 S_{EES}^2 + 2\, b_{EXS}\, b_{EES} \text{ cov (EXS, EES)}\right]^{\frac{1}{2}}$. The last term in the square brackets has been assumed to be zero since we have not computed the covariance term. However, the covariance term and b_{EXS} are both positive and b_{EES} is negative. Thus the figures presented in Table 6.3 are overestimates. Lower bounds on these terms are given by the maximum of $|b_{EXS}|\, S_{EXS}$ and $|b_{EES}|\, S_{EES}$. Substituting the lower bounds reported in footnote 4 of Table 6.3 for the values in the table does not change the ordering reported in the text.

(11) A frequency distribution is said to be positively skewed if there is an elongated tail to the right. This means there are fewer men in the labor force in the upper age groups (say 60-65) than in the lower age groups (say 40-45). Besides being almost intuitive, evidence that this is the case for both men and women is presented in Chapter 8.

(12) This is evidenced by the fact that for the Statistics Canada data, the overall median for the females is 7.7 years of experience whereas the median for the 25-34 age group is 7.8 years. This implies that well over one-half of all the females are less than 35 years of age, and probably over one-half are less than 30 years of age.

CHAPTER 7

THE EFFECTS OF SEX AND MARITAL STATUS ON DISTRIBUTON OF WAGES AND SALARIES(1)

7.1. Introduction

It is well known that the unadjusted average earnings of males in a labor force tend to be higher than those of females of comparable age in the same labor force. The Canadian case for 1970 data is vividly illustrated for five ethnic groups in Chapter 2, Chart 2.4. This chart indicates that the geometric mean earnings of ever-married males are substantially higher than those of females, regardless of marital status. The earnings levels of ever-married males are also higher than for never-married males.

Recently there has been a good deal of research into the sources of the wage differential between males and females. Blinder (1973), Oaxaca (1973), and Brown (1978) have used data for the United States, and Gunderson (1976), Holmes (1976), Robb (1978), and Statistics Canada (Catalogue No. 13-557) have used Canadian data to investigate male-female wage differentials. Each of these studies reached the conclusion that there is pay or wage discrimination against females relative to males. In all these studies, pay discrimination is defined as receiving a wage that is lower than that received by males with comparable qualifications and productivity. In other words, wage discrimination is unequal pay for equal work, and equal pay legislation is designed to prevent this form of discrimination.

A second way that discrimination against females can manifest itself in the labor market is by systematic exclusion of women from higher paying occupations and industries. Equal opportunity legislation is designed to prevent this form of discrimination which is sometimes referred to as employment discrimination. The extent of employment discrimination is difficult to measure since this requires the assumption that equally qualified males and females have the same career motivations and desires. It is not clear that this assumption is satisfied, especially when comparing married males with married females, the two largest groups in the labor force.

See footnote(s) on page 127.

Research results on the pervasiveness of employment discrimination are less clear-cut than for wage discrimination. Oaxaca (1973) has argued that there is significant employment discrimination against women in the United States. Using a different data base and model, Blinder (1973) finds that whereas U.S. males have a superior distribution of employment across occupations, this is more than offset by the greater benefits reaped by females who enter higher paying occupations. Consequently, the combined occupational distribution and reward structure results in net benefits to females. Results reported by Robb (1978) for Canada also provide some evidence of employment discrimination. However, the results of both Blinder and Robb are weakened by virtue of the fact that their procedures bias the occupational distribution results in favor of the males and the occupational benefits results against males.(2) Consequently, their results should be considered as tentative at best.

7.2. The Effect of Marriage on the Intra-Sex Distribution of Wages and Salaries

To a large extent the differential in male-female earnings is the result of females' relatively more tenuous attachment to the labor force for social and biological reasons related to marriage and the bearing and rearing of children. To the extent that this is voluntary supply behavior on the part of women, or to the extent that the differential in earnings results from productivity differences, these earnings differentials cannot be attributed to wage discrimination. If marriage does have important labor supply implications for females, one would expect never-married women to have higher earnings at all ages than do all women. Chart 2.4 supports this hypothesis for all five ethnic groups. The earnings profiles of the never-married women dominate those of ever-married women. These figures also indicate that the earnings of ever-married males are significantly higher than for never-married males at all ages, the exact opposite of that for females. This reversal may, in part, be due to an asymmetric selection process in the "marriage market" as we discuss below.

The usual explanation for the differential between ever-married and never-married males is that ever-married males are more highly motivated than their

---

See footnote(s) on page 127.

never-married counterparts, and this is consistent with the positive coefficients on the dummy variable for being married in the male earnings equations in Chapter 5. Surprisingly, the ever-married coefficients in the female earnings equations reported in Chapter 6 were also positive and significant in spite of the higher gross earnings for never-married females relative to ever-married females. Thus, after controlling for all other factors, ever-married women tend to have higher earnings than do never-married women. In any event, it is clear that marital status must play a crucial role in any investigation of male-female earnings differentials.

One of the principal difficulties in most investigations of male/female earnings differentials is the lack of appropriate data on the labor force ex experience of married women. Of the Canadian studies cited above, by far the best labor force experience data were utilized in the Statistics Canada study. Data on years of actual labor force experience for both males and females were used, but this study is marred by the fact that some cricial factors such as marriage were not taken into account. Robb (1978) controlled for the effect of marriage on female earnings by isolating never-married females 30 years of age and over, and comparing the earnings of that subgroup to the earnings of all males. The reasons for isolating never-married women aged 30 and over is that that group is more likely to be similar to males in terms of career motivation and labor force attachment. Hence, potential labor force experience for never-married females can be approximated by the same method that is used for males.

Our investigation of the effects of sex and marriage on the distribution of earnings is restricted to individuals 30 years of age and over. Any sample of never-married individuals consists of two groups, premarrieds and "confirmed bachelors". Premarrieds are those men and women who intend to marry but have not yet done so, and confirmed bachelors are those who never intend to marry or who once hoped to marry but opportunity passed them by. Premarried women are less likely than confirmed bachelor women to have career motivations similar to those of males. Thus, the female samples are restricted to individuals who are at least 30 years of age in an attempt to eliminate premarried women.

It is important to recognize, however, that if the analysis is restricted restricted to never-married subgroups, the results may not be representative of the general population because of the selection process involved in the marriage market.[3] Marriage, in part, customarily entails an exchange of nonmarket services by the female (childbearing, childrearing, housekeeping, etc.) for the total or partial financial support of the male. Consequently, women who intend to marry are likely to be relatively more "homemaker" oriented and less "career" oriented than confirmed bachelor women. Conversely, men who intend to marry are likely to be relatively more "career" oriented than confirmed bachelor men in order to prepare themselves for the financial obligations of marriage. From the standpoint of the demand for marriage partners, those men who achieve the greatest financial success, holding constant all other factors such as education, attractiveness, intelligence, etc., are the best "catches" as marriage partners. Similarly, those women who achieve the best "homemaker" skills, again holding constant other factors, are in greatest demand as marriage partners. From the supply side, the women who achieve the lowest earnings potential, <u>ceteris paribus</u>, stand to gain the most from marriage since the opportunity costs of foregone wages are relatively small. Thus, <u>ceteris paribus</u>, the confirmed bachelor women are likely to have an average earnings potential which is higher than that for married women. For males, on the other hand, those who achieve the highest earnings potential are more likely to be willing marriage partners since they will be better able to cope with the financial burden of marriage.

Thus, the marriage market is likely to be asymmetric in its selection process. Confirmed bachelor women are likely to have an earnings potential above that of their married counterparts. For males, on the other hand, the reverse is true and married males are likely to have higher earnings potential than are confirmed bachelor males. The raw age-earnings profiles of all five ethnic groups in Chart 2.4 is consistent with this interpretation, but the evidence is weak since there are no adjustments for levels of labor supply, human capital or other attributes. The mean levels of selected attributes of the never-married and ever-married males and females aged 30 and over are presented in Table 7.1. For all five ethnic groups, the geometric mean level of wages and

---

See footnote(s) on page 127.

TABLE 7.1. Average Age, Education, and Geometric Mean Wages and Salaries by Ethnic Group, Marital Status, and Sex(1)

Ethnic group	Marital status	Females: Geometric mean wages	Females: Age	Females: Years of schooling	Females: Weeks worked	Females: Per cent working part-time weeks	Males: Geometric mean wages	Males: Age	Males: Years of schooling	Males: Mean weeks worked	Males: Per cent working part-time weeks
British Group	Never married	4,471	48.5	11.6	46.4	8.8	4,184	44.8	9.7	42.3	12.8
	Ever married	2,211	46.7	10.8	38.7	35.0	7,139	46.9	10.5	46.8	4.8
French	Never married	3,925	45.0	10.1	45.8	12.2	3,825	42.8	8.8	41.3	12.4
	Ever married	2,197	44.6	9.1	38.3	30.1	6,066	44.9	8.6	45.2	6.1
North European	Never married	3,977	44.1	11.5	44.1	12.7	4,565	43.8	9.5	43.7	10.1
	Ever married	2,000	37.5	10.1	37.5	37.4	6,885	45.1	10.1	46.3	5.24
East and South European	Never married	3,954	45.0	10.3	41.6	9.3	4,787	45.3	8.9	44.0	11.8
	Ever married	2,395	45.1	8.7	39.6	30.2	6,750	46.0	9.0	46.0	3.7
Other	Never married	4,034	41.0	11.0	45.5	13.7	4,586	39.4	10.0	41.9	6.8
	Ever married	2,407	42.7	9.2	38.9	26.1	5,678	43.3	9.5	44.0	5.4

(1) Computed for individuals aged 30 years and over.

Source: Public Use Sample Tape.

salaries of the never-married females is substantially higher than their ever-married counterparts. For males the opposite is true; the ever-marrieds have substantially higher wages and salaries than the never-marrieds in every case.

It must be emphasized at this point, the earnings data presented in Table 7.1 and used in the remainder of this chapter includes only wages and salaries since individuals with self-employment earnings are excluded. The reason for excluding them is that self-employment earnings are unlikely to be influenced by sex discrimination(4) in the same way as wages and salaries.

In terms of labor supplied, never-married women worked more weeks and a higher proportion worked full-time weeks than did ever-married women, while just the opposite held for males. This is true for all five ethnic groups. The never-married females had higher education levels than did the ever-married females in all five ethnic groups, while for the males, the ever-marrieds had more education than the never-marrieds in three of the five cases, the exceptions being the French and Other groups. In terms of age, the never-married males were younger than the ever-married males in all cases, and in three of the five cases, never-married females were older than the ever-married females, the exceptions being the French and Other groups.

In summary, then, the never-married females in the labor force are more highly educated and work more weeks than do ever-married women. The result is that the never-marrieds earn substantially higher wages and salaries. For males, the never-marrieds work fewer weeks and earn lower wages and salaries than do ever-married males.

Comparing never-married females with ever-married males, the two groups are of approximately the same age and work approximately the same number of weeks.(5) The never-married females have substantially more years of schooling than their ever-married male counterparts. Offsetting the higher education levels of the never-married females, however, is the fact that a much larger proportion of the females work part-time weeks than do ever-married males. With respect to wages and salaries, however, the ever-married males dominate the never-married females, the latter group earning only from 58% to 71% of the former.

See footnote(s) on page 127.

Two earnings functions were estimated for the never-married and ever-married subgroups of both males and females in each ethnic group. The two earnings functions differ in the sense that one of the specifications, the full model, includes controls for occupation and industry of employment whereas the reduced model does not.(6) Within each ethnic group, statistical tests were performed to test the hypotheses that the coefficients did not differ between each possible pairwise sex-marital status comparison (e.g., never-married males verus versus ever-married males, never-married males versus never-married females, etc.) These tests were conducted for both the full and reduced models making a total of 12 pairwise comparisons for each ethnic group. Of all these comparisons within each ethnic group, the hypothesis that there is no difference between any pair was rejected at the 1% level in every case except for the comparisons of never-married females versus ever-married females, and never-married females versus never-married males. For the never-married/ever-married female comparison, the hypothesis that there is no difference in the coefficients can be rejected at the 5% significance level only for the British, full and reduced model, and for the British full model, the hypothesis of no difference cannot be rejected at the 1% level. For the remaining ethnic groups there does not seem to be any significant difference in the estimated coefficients. Thus, with the exception of the British group, there does not appear to be any significant difference in the pay structure between never-married and ever-married women. Concerning the never-married female/never-married male comparisons, the hypothesis of no difference could not be rejected at the 5% level for the Other ethnic groups, and for the North European group, the hypotheses of no difference in the reduced model could be rejected at the 5% level but not at the 1% level.

Thus, it appears that the coefficient estimates, and hence pay structures differ among sex-marital status groups within ethnic groups. The only important deviation from this pattern is the never-married/ever-married female pattern. The results indicate that, with the possible exception of the British group, there is no statistically significant difference in the reward structure between ever-married and never-married females.

The return to schooling by marital status, sex, and ethnic group are

---

See footnote(s) on page 127.

presented in Table 7.2. The results indicate that the return to schooling is higher in the reduced model than in the full model in every case, but the differences are far more pronounced for the females than for the males.(7) Thus, for both ever-married and never-married females, higher levels of schooling appear to be positively associated with employment in higher paying occupations and industries. Controlling for occupation and industry of employment in the full model results in substantially lower estimated returns to schooling relative to the reduced model since, in the latter, the earnings increase associated with the occupational/industrial employment shift is picked up by the schooling variable. This effect is much less pronounced for the males than it is for the females, which suggests that the role of education as a screening device for employment in higher paying occupations and industries is less important for males than for females.

The gross earnings differentials between never-married and ever-married individuals by sex and ethnic group are partitioned in Table 7.3. (The method of partitioning is the same as that used in Chapter 6 and discussed in Appendix H.) The gross differential is the percentage by which the geometric mean wages and salaries of the never-marrieds exceed the geometric mean wages and salaries of ever-marrieds belonging to the same group. This differential is partitioned into two parts: the first is that part which is associated with or explained by differences in attribute levels, while the second is the part which arises because of differences in rates-of-return to the attributes (i.e., different pay structures between the groups).

The figures for females indicate that the major portion of the wage and salary differential between the never-married and ever-married groups can be explained by differences in attribute levels rather than differences in rates-of-return to those attributes. This comes as no surprise since the rates-of-returns on attributes (i.e., the coefficients) were not significantly different, in the statistical sense, between the ever-married and never-married females except for the British group. Even for this group, however, the estimates indicate that of the 70.4% gross differential, the major portion (between 56.4% and 63.3%) is associated with attribute level differences, and only a small

---

See footnote(s) on page 127.

TABLE 7.2. Return to Schooling by Marital Status, Sex, and Ethnic Group(1)

Ethnic group	Marital status	Female Reduced	Female Full	Male Reduced	Male Full
British	Never married	7.03	3.68	2.91	2.50
	Ever married	5.40	2.86	3.72	3.03
French	Never married	6.69	2.09	2.95	2.88
	Ever married	5.00	2.14	2.86	2.55
North European	Never married	2.56	1.81	0.83	0.16
	Ever married	4.44	2.03	1.65	1.58
East and South European	Never married	4.73	2.04	4.95	4.02
	Ever married	3.42	2.02	2.28	1.63
Other	Never married	5.82	5.23	4.31	3.03
	Ever married	3.54	1.64	2.78	2.40
Simple average		4.86	2.55	2.87	2.38

(1) Computed for individuals aged 30 years and over and not reporting earnings from self-employment.

Source: Public Use Sample Tape.

TABLE 7.3. Partitioning of Wage and Salary Differences between Never-married and Ever-married Individuals 30 Years and Over by Ethnic Group and Sex(1)

Ethnic group	Gross pay differential	Attribute level effects				Rate-of-return effects			
		Reduced model		Full model		Reduced model		Full model	
		Lower bound	Upper bound	Lower bound	Upper bound	Lower bound	Upper bound	Lower bound	Upper bound
A. Never-married versus ever-married females									
British	70.4	56.4	62.7	58.8	63.3	7.8	14.1	7.2	11.8
French	58.0	40.1	51.9	42.2	54.3	6.2	17.9	3.6	15.7
North European	68.7	62.4	67.5	71.4	71.6	1.1	6.3	- 2.7	- 3.0
East and South European	50.1	46.8	53.7	47.5	53.7	3.2	- 3.6	2.4	- 3.8
Other	51.6	43.5	52.5	53.0	55.6	- 0.9	8.0	- 1.5	- 4.1
B. Never-married versus ever-married males									
British	-53.4	-21.4	-21.8	-24.5	-26.3	-32.8	-33.1	-28.1	-30.0
French	-46.1	-11.3	-13.8	-14.4	-16.0	-32.4	-34.8	-29.6	-31.2
North European	-41.0	-11.1	-11.2	-13.1	-16.7	-30.0	-30.0	-24.7	-28.3
East and South European	-38.1	- 6.5	-11.8	- 9.5	-13.4	-22.7	-27.9	-21.6	-25.5
Other	-21.3	- 4.7	- 5.2	- 6.2	-10.0	-14.1	-14.6	-11.1	-14.8

(1) The numbers are percentage differences in geometric mean wages and salaries and are computed as 100 times the differences between means of logarithms. All differences are computed as never-married minus ever-married. Thus, a negative number indicates an advantage to the ever-married group whereas a positive number indicates an advantage to the never-married group.

Source: Public Use Sample Tape.

part of the total differential arises because of the differences in pay structure (rates-of-return effects). Thus, at least 80% of the gross differential is associated with differences in attribute levels such as higher levels of education, more weeks and hours worked, etc., and not more than 20% of the gross differential is related to differing pay structures.

The pattern is the same for the other ethnic groups. The major portion of the gross differential is associated with attribute levels and very little results from differing rates-of-return.

For males, the situation is completely reversed. The negative gross differential indicates that never-married men earn lower geometric mean wages and salaries than do ever-married men, the latter earning from 21.3% more for the Other group, to 53.4% more for the British group. For each of these ethnic groups, the amount of the never-married/ever-married differential associated with attribute level differences is smaller (in absolute value) than that part associated with differing pay structures. Thus, differences in rates-of-return to attributes are a more important source of the wage and salary level differentials between ever-married and never-married males than are the differences in levels of attributes between these two groups.

The earnings level differences associated with rate-of-return differences have frequently been labeled as discrimination between the groups being studied (Blinder, 1973; Oaxaca, 1973; Masters, 1974; Holmes 1976; Robb, 1978). However, these studies have usually involved comparisons between the sexes or racial groups where there is a good deal of popular belief that wage discrimination does exist. To our knowledge, no one has ever suggested that there is wage discrimination against never-married males relative to ever-married males. The higher earnings of ever-married relative to never-married males are usually attributed to motivational differences between these groups: the married male works harder than the never-married male because the former is frequently the principal bread winner for wife and family. If that is in fact the case, then the differences in the coefficient estimates for the never-married and ever-married earnings functions arise because of "motivational differences" which are not adequately captured in the model specification. However, it is these coefficient differences which give rise to the "residuals" which have often been labeled as discrimination.

Given the weak corroborating evidence of pay discrimination against never-married males relative to ever-married males, it is probably better to interpret the rate-of-return effects as primarily reflecting motivational differences that arise from the factors associated with marriage (such as the mate selection process). This being the case, it is only natural to raise questions concerning the meaning of the large residual between races and between sexes identified as discrimination in other studies. It is not possible that these "residuals" also reflect other factors besides pay discrimination which were not taken into account in the specification?

In summary, the wage and salary differential between never-married and ever-married females can largely be explained by differences in the levels of various attributes between these two groups, and only a minor portion of the differential is due to rates-of-return differences between the two groups. For males, however, differences in rates-of-return to attributes account for over half of the gross wage and salary differential between never-married and ever-married males.

Wage and salary differentials between males and females are analyzed in the next section.

7.3. The Effect of Sex on the Distribution of Wages and Salaries

One the basis of the selected attribute levels discussed in connection with Table 7.1, it is clear that never-married females are much more similar to males in terms of work effort than are ever-married females. In terms of weeks worked, never-married females are very similar to ever-married males, but in terms of the proportion working part-time rather than full-time weeks, never-married females resemble never-married males more closely than they do ever-married males. The geometric mean wage and salary levels of the male and female never-married groups are also very similar. Because of the similarity of the attributes of the never-married females to those of males in general, the wage and salary differentials between never-married females on the one hand, and both never-married and ever-married males on the other, are discussed in this section. For comparison ever-married females are briefly discussed relative to all males.

The pay differentials between never-married males and females for the five ethnic groups are partitioned in the first section of Table 7.4. The gross differentials indicate that the wages and salaries of never-married females in the British and French ethnic groups slightly exceed those of never-married males by 6.6.% and 2.6%, respectively. For the remaining three ethnic groups, however, the pay level of the never-married males exceeds that of their never-married female counterparts. Comparing the signs of the attribute level and rate-of-return effects, it is apparent that these effects work in opposite directions. For every ethnic group, the never-married females have a relatively more desirable attribute package than do never-married males. With respect to the rates-of-return effects, however, the females are at a disadvantage. The rates-of-return advantage for the males more than offsets the attribute level advantage for three of the five groups, with the net result that the never-married males in those groups earn more than the never-married females. For the British and French groups, however, the two effects nearly offset each other with the result that the never-married males and females have approximately the same pay level. The pattern is the same for both the full and reduced model estimates.

The wage and salary differentials between never-married females relative to ever-married males are partitioned in the second section of Table 7.4. The gross differentials indicate that ever-married males earn from 34.1% to 57.2% more than never-married females in the same ethnic group. The attribute level effects are negligible in most cases indicating that never-married females and ever-married females have very similar attribute packages. In fact, the lower and upper bounds in the full model specification include zero in the estimated range of the attribute level effects for all of the ethnic groups. Thus, it is not clear whether the ever-married males or the never-married females have an attribute level advantage.

TABLE 7.4. Partitioning of Wage and Salary Difference Between Males and Females 30 Years and Over by Ethnic Group and Marital Status(1)

		Attribute level effects				Rate-of-return effects			
		Reduced model		Full model		Reduced model		Full model	
Ethnic group	Gross pay differential	Lower bound	Upper bound	Lower bound	Upper bound	Lower bound	Upper bound	Lower bound	Upper bound
A. Never-married males versus never-married females									
British	- 6.6	-26.4	-32.3	-19.8	-33.8	19.7	25.7	13.9	27.9
French	- 2.6	-18.8	-22.9	-14.1	-24.6	16.2	20.3	12.7	23.2
North European	13.8	- 3.6	-15.3	- 3.4	-38.2	17.4	29.0	17.7	52.5
East and South European	19.1	-20.7	-25.0	-14.7	-37.6	39.8	44.1	33.9	56.8
Other	12.8	-10.2	-18.0	-14.7	-60.6	23.0	30.8	29.4	75.3
B. Ever-married males versus never-married females									
British	46.8	- 3.0	- 4.6	- 5.8	6.8	50.9	52.5	41.7	54.3
French	43.5	- 5.6	- 9.3	2.5	- 7.5	49.2	52.9	41.7	51.6
North European	54.8	0.3	4.4	- 4.8	14.7	50.5	54.7	41.0	60.5
East and South European	57.2	- 6.9	- 7.6	3.1	-22.2	60.4	61.1	57.2	76.4
Other	34.1	- 4.6	-12.4	2.2	-56.5	38.9	46.7	33.6	92.4

See footnote(s) at end of table.

TABLE 7.4. Partitioning of Wage and Salary Difference Between Males and Females 30 Years and Over by Ethnic Group and Marital Status(1) (Concluded)

		Attribute level effects				Rate-of-return effects			
		Reduced model		Full model		Reduced model		Full model	
Ethnic group	Gross pay differential	Lower bound	Upper bound	Lower bound	Upper bound	Lower bound	Upper bound	Lower bound	Upper bound
C. All males versus ever-married females									
British	113.	41.6	56.8	52.5	56.6	57.6	72.8	58.4	62.5
French	97.	25.8	42.6	37.1	43.3	54.7	71.5	54.6	60.8
North European	121.	45.7	69.2	58.9	74.0	51.4	74.9	48.6	63.8
East and South European	101.	29.0	42.5	38.9	49.3	58.6	72.2	52.4	62.9
Other	84.	23.3	40.5	29.6	33.1	43.2	60.4	52.2	55.7

(1) The numbers are percentage differences in geometric mean wages and salaries and are computed as the differences between means of logarithms multiplied by 100. All differences are computed as male minus female. Thus, positive numbers indicate advantages to males whereas negative numbers indicate advantages to females.

Source: Public Use Sample Tape.

With respect to the rate-of-return effects, however, the ever-married males have a distinct advantage relative to never-married females. The magnitudes are substantial and indicate that the ever-married males receive much higher rates of return on their attributes than do never-married females. For example, the lower and upper bound estimates based on the reduced model for the British group indicate that ever-married males receive between 50.9% and 52.5% more pay than do never-married British females because the males have higher rates-of-rates on their attributes. The comparable estimates for the full model indicate that the never-married males receive from 41.7% to 54.3% more pay than never-married females because of the higher rates-of-return. The magnitudes for the remaining four ethnic groups are similar.

The wage and salary differentials between ever-married females and all males are partitioned in the third section of Table 7.4. The gross differentials indicate that the geometric mean wages and salaries of males are approximately double those of ever-married females. The differential resulting from rates-of-return differences is approximately the same for the ever-married female/all male comparison as it is for the never-married female/ever-married male comparison. The big difference between these two comparisons is in the attribute level effects. While the never-married females and ever-married males had very similar attribute levels, it is clear that ever-married females have a much less desirable attribute package than do all males.

The substantial rates-of-return effects in Table 7.4 have been labeled by some researchers as discrimination. Care must be used in such an interpretation, however. As discussed above in connection with the ever-married/never-married male comparison, the difference in rates-of-return between two groups may merely be a reflection of factors that have not been controlled for in the analysis. For example, suppose never-married females are equivalent to ever-married males in terms of all the factors considered in the analysis, but ever-married males are more highly motivated and, consequently, work harder than the never-married females. If the labor market rewards the males for the resulting higher productivity, the only way this can be reflected in the analysis is through higher rates-of-return on the attributes considered in the analysis. It is this type of bias that may account for a portion of the ever-married/never-married male rates-of-return differentials. If that is the case, it would not be correct to label the rates-of-return effects as wage discrimination. At this time, it is open to question as to whether the male/female rates-of-return differentials should be labeled as discrimination.

Some people would argue that the reason females aren't as highly motivated as males (if that is, in fact, the case) is because of the typical social and biological role of women in society. As a consequence of this stereotyping, the female is viewed as someone who will only be in the labor force intermittently or temporarily. As a result, women receive lower pay and have fewer opportunities for advancement than do men in comparable situations. Thus, there is little incentive for women to work hard at their jobs or to acquire skills that are rewarded in the labor market. This disincentive can be considered discrimination in a larger sense.

The never-married women considered in this study, however, certainly invested time and effort acquiring education. On this score, they compared very favorably with the males. They also compared favorably in terms of weeks of work, and these factors are prima facie evidence of motivation. There is a good deal of popular belief that pay discrimination against women exists and other researchers have labeled rates-of-return differentials such as we have found as pay discrimination. All of this suggests that at least a portion of the large rates-of-return effects between ever-married males and never-married females could be labeled discrimination. However, if the male-female rates-of-return differentials are labeled as discrimination, then the rates-of-return differentials between never-married and ever-married males discussed in the preceding section should also be labeled as discrimination.

7.4. The Effect of the Distribution of Employment on Wage and Salaries

A second way that labor market discrimination against a group can manifest itself is through employment discrimination, which is defined as the systematic exclusion of the disadvantaged group from higher paying occupations and industries. Within the current framework of analysis, employment discrimination would be reflected by an increase in the absolute size of the attribute-level effects when controlling for the industry and occupation of employment.[8] Thus if employment discrimination is present, the absolute magnitude of the attribute level effects computed on the basis of the full model should be larger than those of the reduced model. Since the attribute level effects plus the rate-of-return effects for each model (full and reduced) must sum to the gross differential, there must be an offsetting change in the rate-of-return effects. In particular, employment discrimination would be reflected by a reduction in the absolute size of the rate-of-return effects in the full model relative to the reduced model.

The evidence presented in Tables 7.3 does not provide any support for the contention that there is employment discrimination between never-married and ever-married individuals of the same sex. Similarly, the evidence in Table 7.4 provides no support for the hypothesis that never-married females are discriminated against by being systematically excluded from higher paying occupations and industries.

This finding is in direct contrast to the conclusion reached by Oaxaca (1973, p. 708) for women in the United States. He argued that it was the concentration of women in lower paying jobs that produces large pay differentials between males and females, and that this effect is far more important than unequal pay for equal work. One possible reason for the differences in the conclusions reached in Oaxaca's study and the conclusion reached in this study is that Oaxaca used a somewhat more detailed set of industries and occupations than were used in the present analysis.[9] It is possible that employment discrimination would be indicated if more detailed data were utilized on occupation and industry of employment.

See footnote(s) on page 127.

FOOTNOTES

(1) Richard P. Beilock is a co-author of this chapter.

(2) The occupational distribution advantage for males was calculated as $\sum_i b_i^m (\bar{X}_i^m - \bar{X}_i^f)$, and the occupational benefits advantage for males was calculated as $_i\ \bar{X}_i^f(b_i^m - b_i^f)$, where $\bar{X}_i^m$ and $\bar{X}_i^f$ indicate the proportion of males and females in occupation i and b_i^m and b_i^f are the male and female returns or benefits to being employed in occupation i relative to some reference occupation, and the summation is over all occupations included in the model specification. If males and females respond to differential benefits in occupations, they should do so by entering those occupations with relatively high benefits. Thus, if b_i^m is high relative to b_i^f for some particular occupation, then that occupation is relatively more attractive for males than for females and $\bar{X}_i^m$ should be large relative to $\bar{X}_i^f$. The result is that weighting the intersex occupational distribution of employment $(\bar{X}_i^m - \bar{X}_i^f)$ by the male benefits biases the result in favor of males. Similarly, weighting the intersex differentials within occupations $(b_i^m - b_i^f)$ by the female occupational distribution biases the result in favor of females (against males). The solution to this index number problem is to evaluate each of these terms twice, once using male weights and once using female weights (see Appendix H).

(3) The following argument is based on the theory of marriage by Becker (1976).

(4) The earnings data underlying Chart 2.4 includes individuals with self-employment income as well as wages and salaries.

(5) Concerning age, the never-married females and ever-married males were virtually identical in age for the French, North European and East and South European groups, whereas the females were somewhat older for the British group and the males were older for the Other group. The never-married females worked almost the same number of weeks as the ever-married males of the British and French groups. Ever-married males of the North European and East and South European groups worked more weeks than the

never-married females, while just the opposite was true for the Other group.

(6) The specifications of both the full and reduced model are discussed in Appendix I. Due to the large number of regressions involved, the estimates are not reported.

(7) For females, the simple average across all ethnic and marital-status groups increases from 2.55 in the full model to 4.86 in the reduced model, whereas the comparable figures for the males are 2.38 and 2.87 for the full and reduced male models, respectively. The male figures exhibit a much smaller difference between the two models.

(8) The reason for this is as follows. Let $\bar{X}_i^1$ and $\bar{X}_i^2$ denote the proportion of the workers in Groups 1 and 2 that are employed in industry i. (That is, $\bar{X}_i^1$ and $\bar{X}_i^2$ are the means of the ith occupational dummy variable for Groups 1 and 2). If occupation i is a high paying occupation relative to the excluded category, then the estimated coefficients, b_i^1 and b_i^2, should be positive. If both groups have equal desires to work in occupations with high pay, then $\bar{X}_i^1$ and $\bar{X}_i^2$ should be approximately equal and the attribute level effects for that occupation, $b_i^1\ (\bar{X}_i^1 - \bar{X}_i^2)$ or $b_i^2\ (\bar{X}_i^1 - \bar{X}_i^2)$, will be negligible. On the other hand, if Group 2 is systematically excluded from higher paying occupations, or if individuals in this group have less desire to work in those occupations than do individuals in Group 1, then $b_i^1\ (\bar{X}_i^1 - \bar{X}_i^2)$ and $b_i^2\ (\bar{X}_i^1 - \bar{X}_i^2)$ will be positive. For occupations with low pay relative to the excluded category, b_i^1 and b_i^2 will be negative. If Group 2 is systematically excluded from higher paying jobs, or the individuals in this group do not want to provide the effort necessary to work in those jobs, then they must necessarily be concentrated in low paying occupations and $(\bar{X}_i^1 - \bar{X}_i^2)$ will be negative for those occupations. Thus, summing over all occupations, $\sum_i b_i^1\ (\bar{X}_i^1 - \bar{X}_i^2)$ and $\sum_i b_i^2(\bar{X}_i^1 - \bar{X}_i^2)$ will be positive if the distribution of employment has an important effect on increasing the level of wages and salaries in favor of Group 1.

(9) Oaxaca utilized a total of 10 occupations and 16 industries. In the present analysis, only 11 occupations and nine industries were utilized.

CHAPTER 8

THE DISTRIBUTION OF HOURS AND WEEKS OF WORK

One of the major shortcomings of our analytical models is their failure to explain the number of weeks individuals work. The variations in weeks worked is a important source of earnings inequality for both males and females. The inequality analysis in Chapter 5 indicated that over 19% of the variance in the logarithm of male earnings is attributed to variations in weeks worked. The analysis of female earnings in Chapter 6 showed that variations in weeks worked were by far the largest source of earnings inequality for each of the female ethnic groupings analyzed.

This chapter examines distributions of the number of weeks worked by sex, age, and ethnic groups. While this does not explain weeks worked, it does provide information on how the number of weeks worked varies across age categories and ethnic groups. Since the number of hours worked per week is the other dimension of labor supply that is important in explaining individual earnings, the chapter includes a brief discussion of the distributions of the hours of work by sex.

It was necessary to exclude individuals who did not work in 1970 from the samples used in the econometric analyses because these individuals had zero earnings. Consequently, the dependent variables are undefined for those observations. However, it is informative to describe some of the characteristics of those individuals who did not work. Hence, the distributions analyzed in this chapter include observations on all individuals over 14 years of age in the provincial files of the Public Use Sample. This is a larger sample than that used in the analysis presented in Chapter 5, 6, and 7.

In Table 8.1. the distributions of weeks of work by broad age categories are presented for both males and females. Comparing the overall distributions of weeks of work for males and females (column totals for males and females), it can be seen that 18.7% of the males in the sample did not work in 1970 compared with 54.3% of the females. At the other end of the weeks worked spectrum, 49.6% of the males worked full-years (i.e., in excess of 48 weeks) compared with only 20.9% of females. While the number of observations at the ends of the spectrum for males

TABLE 8.1. Distribution of Males and Females by the Number of Weeks Worked and by Age.

Age groups	Number of weeks worked(1)					Row total(2)
	0	1-13	14-39	40-48	49-52	
			Males			
15-24	6,241 32.1	3,299 17.0	4,191 21.6	1,578 8.1	4,119 21.2	19,428 26.7
25-44	931 3.5	662 2.5	3,451 2.9	3,435 12.9	18,245 68.3	26,724 36.7
45-64	1,620 8.3	539 2.8	2,267 11.6	2,466 12.7	12,578 64.6	19,471 26.7
65 and over	4,866 67.1	297 4.1	583 8.0	322 4.4	1,183 16.3	7,251 10.0
Column total	13,658 18.7	4,797 6.6	10,492 14.4	7,801 10.7	36,126 49.6	72,874 100.0
			Females			
15-24	8,607 44.5	2,596 13.4	3,266 16.9	1,283 6.6	3,568 18.5	19,320 26.1
25-44	13,042 49.5	1,686 6.4	3,248 12.3	1,904 7.2	6,493 24.6	26.373 35.6
45-64	10,723 54.6	807 4.1	1,727 8.8	1,414 7.2	4,983 25.4	19,654 26.6
65 and over	7,785 89.8	101 1.2	235 2.7	118 1.4	431 5.0	8,670 11.7
Column total	40,157 54.3	5,190 7.0	8,476 11.5	4,719 6.4	15,475 20.9	74,017 100.0

(1) The first entry in each cell is the number of observations in the cell and the second entry is the number of observations as a percent of the row total.

(2) The second entry is the row total as a percent of the overall total.

Source: Public Use Sample Tape.

and females are radically different, between the endpoints the distributions are fairly similar: 6.6% of males versus 7.0% of females worked from 1-13 weeks; 14.4% of males versus 11.5% of females worked 14-39 weeks; and 10.7% of males versus 6.4% of females worked 40-48 weeks.

Comparing the distributions within age groups, the male and female distributions in the 15-24 age groups are probably the most similar of any of the age groups. The reasons for this may be that the samples on which these distributions are based include full- and part-time students, and the majority of never-married individuals appear in this age group as well. At any rate, 32.1% of males versus 44.5% of females in this age group did not work, and 21.2% of males versus 18.5% of females worked full-years. The distributions of males in the 25-44 age group and 45-64 age group are very similar. The same is true for females, but the distributions are radically different between the males and females. For the males, only 3.5% and 8.3% in the 25-44 and 45-64 age groups respectively did not work, whereas the comparable figure for females are 49.5% and 54.6%. Thus, while approximately one-half of the females in these age groups did not work during 1970, 94.5% of the males worked at least part-years, and 66.7% worked full-years. On the other hand only 24.9% of the females in these age groups worked full-years. For the 65 and over age group, 89.8% of the females and 67.1% of the males did not work. Only 5% of females in this age group worked full-years compared with 16.3% of males. Thus, for every age group, a higher proportion of males worked full-years than females, and a higher proportion of females than males did not work. To what extent these differences are voluntary supply effects (e.g., women withdrawing from the labor force for childrearing and homemaking), and to what extent they are involuntary demand effects (greater difficulty experienced by females in finding jobs) we have not been able to determine. Nevertheless, whatever the cause, these differences are important in explaining the differences in the distribution of earnings between males and females.

8.1. Weeks of Work by Ethnic Group

The distributions of the numbers of males and females by weeks of work for different ethnic groups are presented in Table 8.2. The ethnic groups in this table correspond to those used in Chapter 5 for males and those used for females in Chapters 6 and 7.

TABLE 8.2. Distribution of Individuals by the Number of Weeks Worked, Ethnic Group, and Sex

Ethnic group	Number of weeks worked(1) 0	1-13	14-39	40-48	49-52	Row total(2)
			Males			
British	5,856 18.3	2,128 6.6	4,399 13.7	3,207 10.0	16,443 51.3	32,033 44.0
French	4,471 22.0	1,362 6.7	3,017 14.8	2,087 10.3	9,385 46.2	20,322 27.9
North European	1,186 13.8	524 6.1	1,239 14.4	990 11.5	4,654 54.2	8,593 11.8
East European	709 19.0	235 6.3	515 13.8	411 11.0	1,857 49.8	3,727 5.1
Italian	321 12.4	139 5.4	441 17.0	420 16.2	1,267 48.9	2,588 3.6
Jewish	193 16.6	94 8.1	136 11.7	120 10.4	618 53.2	1,161 1.6
Oriental	131 22.3	38 6.5	90 15.3	69 11.8	259 44.1	587 0.8
Native Indian	278 36.6	103 13.6	166 21.8	69 9.1	144 18.9	760 1.0
Other	513 16.5	174 5.6	489 15.8	428 13.8	1,499 48.3	3,103 4.3
Male total	13,658 18.7	4,797 6.6	10,492 14.4	7,801 10.7	36,126 49.6	72,874 100.0
			Females			
British	17,697 52.3	2,565 7.6	4,033 11.9	2,127 6.3	7,401 21.9	33,823 45.7
French	12,826 61.3	1,169 5.6	1,976 9.4	1,148 5.5	3,804 18.2	20,923 28.3
North European	4,049 50.0	686 8.5	1,015 12.5	556 6.9	1,791 22.1	8,097 10.9
East and South European	3,496 59.5	498 7.0	905 12.8	566 8.0	1,603 22.7	7,068 9.5
Other	2,089 50.9	272 6.6	547 13.3	322 7.8	877 21.4	4,106 5.5
Female totals	40,157 54.3	5,190 7.0	8,476 11.5	4,719 6.4	15,475 20.9	74,017 100.0

(1) The first entry in each cell is the number of males or females in the cell. The second entry is the number of individuals as a percent of the number in the ethnic group.

(2) The second entry in each cell in this column is the number of individuals in the ethnic group as a percent of the total number of males or females.

(3) For males, this group combines the Negro/West Indian with the Other and unknown group of Chapter 5.

Source: Public Use Sample Tape.

As might be expected, there is great variability among the ethnic groups in the distribution of the weeks of work. Whereas only 18.7% of all males did not work in 1970, nearly twice that percentage (36.6%) of Native Indians did not work at all, and over one-half of male Native Indians over 14 years of age either did not work or worked less than 14 weeks. This contrasts sharply with the Oriental group which had the next highest percentage of males who did not work (22.3%). The ethnic groups with the lowest percentages of males who did not work were the Italians and Jews with 12.4% and 16.6% respectively.

At the other end of the spectrum, 49.6% of all males worked full-years (in excess of 48 weeks), but only 18.8% of male Native Indians worked full-years. This is far lower than the 44.1% of the Oriental group which had the second lowest percentage of full-time workers. The North European and Jewish groups had the highest proportion of males working full-time with 54.2% and 53.2% respectively.

Distributions of weeks of work by ethnic group for females show much less variablity among ethnic groups than do the distributions for males. To a large extent, this is due to the fact that the numerically smaller ethnic groups have been aggregated.(1) Of all females 54.3% did not work. In this regard, the French group had the largest percentage who did not work (61.3%), and this was substantially higher than any of the other groups. While 20.9% of all females worked full-years, only 18.2% of the French group worked full-years. Thus, the French group had both the largest proportion of females who did not work, and also the lowest proportion who worked full-years. The French group was the only female group that differed substantially from the overall pattern for females.

The dramatic differences in the distributions of the numbers of weeks worked are clearly important in explaining the differences in earnings among ethnic groups discussed in Chapter 2 for both males and females (Charts 2.2 and 2.3), and discussed further in Chapters 5, 6, and 7.

8.2. Hours of Work by Sex

The distributions of hours usually worked per week by sex are given in Table 8.3. It was noted in Chapter 4 that these data do not necessarily refer to the individual's labor force activity in 1970. Rather, these data relate to the

See footnote(s) on page 137.

job held in the week prior to enumeration, or to the job of longest duration since January 1, 1970, if the individual was not employed in the enumeration week. Nevertheless, the data in Table 8.3 provide some information on the distribution of hours of work. As might be expected, the number of males and females who reported zero hours of work are lower than the number who reported zero weeks of work in 1970. This is because the question on hours of work related not only to jobs held in 1970, but also to jobs that may been held in 1971 prior to enumeration.

TABLE 8.3. Distribution of Individuals by Hours Usually Worked per Week and by Sex(1)

	Hours usually worked								
Sex	0	1-19	20-29	30-34	35-39	40-44	45-49	50 and over	Row total
Male	12,435	3,345	1,401	1,355	7,533	28,691	7,152	10,962	72,874
	17.1	4.6	1.9	1.9	10.3	39.4	9.8	15.0	100.0
Female	37,749	6,170	3,231	2,138	8,686	12,379	1,655	2,009	74,017
	51.0	8.3	4.4	2.9	11.7	16.7	2.2	2.7	100.0

(1)The first entry in each cell is the number of individuals and the second entry is the number of individuals as a per cent of the total for the sex.

Source: Public Use Sample Tape.

The 51.0% of females reporting zero hours of work is much larger than the 17.1% of males reporting zero hours. Whereas 64.2% of all males worked 40 or more hours per week, only 21.6% of females worked 40 or more hours per week. Considering only individuals who had nonzero hours of work, 67.4% of males worked 40 or more hours per week compared with 44.2% of females. These substantial differences in the distribution of hours of work between males and females will tend to result in lower earnings for females than males. This is consistent with the observations made on the relative earnings of males and females in Chapters 2 and 7.

8.3. Characteristics of Nonworking Males and Females

This study has been primarily concerned with the distribution of earnings. However, 18.7% of males and 54.3% of females over 14 years of age did not work in 1970 and consequently had no earnings. Table 8.4 provides distributions of non-working males and females in broad age groupings by other family income. Of the

TABLE 8.4. Distribution of Non-working Males and Females by Age and Other Family Income.(1)

Age	Other family income(2) 0	1-2,499	2,500-8,499	8,500 and over	Row total(3)
			Males		
15-24	51	751	2,073	3,366	6,241
	0.8	12.0	33.2	53.9	45.7
25-44	33	501	293	104	931
	3.5	53.8	31.5	11.2	6.8
45-64	89	788	556	187	1,620
	5.5	48.6	34.3	11.6	11.9
65 and over	78	2,082	2,225	481	4,866
	1.6	42.8	45.7	9.9	35.6
Column total	251	4,122	5,147	4,138	13,658
	1.8	30.2	37.7	30.3	100.0
			Females		
15-24	76	1,125	3,258	4,148	8,607
	0.9	13.1	37.9	48.2	21.4
25-44	113	1,245	5,046	6,638	13,042
	0.9	9.5	38.7	50.9	32.5
45-64	159	2,200	3,615	4,749	10,723
	1.5	20.5	33.7	44.3	26.7
65 and over	153	4,478	2,339	815	7,785
	2.0	57.5	30.0	10.5	19.4
Column total	501	9,048	14,258	16,350	40,157
	1.2	22.5	35.5	40.7	100.0

(1) Other family income is defined as total family income minus the individual's employment income.

(2) The first entry in each cell is the number of observations in the cell, and the second entry is the number of observations as a percent of the number of males or females in the age group (row total).

(3) The second entry is the number of males or females in the age group as a percent of the total number of males or females in the sample.

Source: Public Use Sample Tape.

individuals in the sample who did not work in 1970, 1.8% of the males and 1.2% of the females reported zero other family income. (These individuals presumably lived on savings.) A further 30.2% of males and 22.5% of females reported other family incomes of less than $2,500 and 69.7% of males and 59.3% of females had other income of less than $8,500.

Concerning the age distribution of nonworkers, the majority of males were under 25 or over 64 years of age (45.7% and 35.6%, respectively), and only 18.7% of nonworking males were in the 25-64 age bracket. This contrasts sharply with the distribution for females where 59.2% were in the 25-64 age bracket and only 21.4% were under 25 and 19.4% were over 64. This large difference in the age distributions between nonworking males and females is no doubt largely due to the high proportion of mothers and homemakers among the females in the 25-64 year age group.

The distributions of nonworkers by other family income within age groups vary sharply among different age groups. In the 15-24 age group, 53.9% of males and 48.2% of females had other family income of $8,500 or over, whereas only 9.9% of males and 10.5% of females 65 and over had other family incomes of that magnitude. In the middle age groups, 57.3% and 54.1% of nonworking males in the 25-44 and 45-64 age groups reported other family incomes of less than $2,500, contrasting sharply with the 10.4% and 22.0% of non-working females in these age-income groups. Some 50.9% and 44.3% of the women in the 25-44 and 45-64 age groups reported other family incomes in excess of $8,500, whereas only 11.2% and 11.6% of the nonworking males in these age groups had other family incomes of that magnitude.

The nonworking females in the middle age groups appear to have higher family incomes than males in these age groups. However, a slightly larger proportion of nonworking males than females in the 15-24 age group reported family incomes of $8,500 or over, and a larger proportion of nonworking females than males in the 65 and over age group reported other family incomes of less than $2,500. However, looking at the distributions for all nonworking males and females, the females appear to be "better off" than males in the sense that a higher proportion of females have other family incomes of $8,500 or over, and a lower proportion of females report other family incomes of less than $2,500.

FOOTNOTE

(1) Although the results are not presented, the distributions of weeks of work corresponding to the ethnic groups delineated for males were also tabulated. The variability among the larger number of female ethnic groups was nearly as great as that for the males.

APPENDIX A

DETERMINATION OF THE RATE-OF-RETURN TO HUMAN CAPITAL

In this appendix we discuss the determination of the rate-of-return to human capital for an individual. This amounts to a digression on the optimal accumulation of human capital.

Becker (1967, 1975) has discussed the optimal investment in human capital by individuals within a framework of the demand for and supply of investment funds.[(1)] With perfect information in a world of certainty, a rational individual should continue to invest in human capital in every period, until the internal rate-of-return from additional investment is equal to the opportunity cost (rate of interest) on additional investment funds (where the costs and returns include psychic as well as monetary components). The appropriate internal rate-of-return is that which equates the present value of all increases in future income resulting from the investment to the discounted cost of the investment.

In any particular time period, the production of human capital (h) is assumed to require inputs of previously accumulated human capital (H), the share of the investor's time devoted to the accumulation of human capital (T), and the purchased market resources (M). Formally,

$$h = G(M, T, H, B), \qquad (A.1)$$

where B is a vector denoting the individual's ability to use previously accumulated human capital to earn income in the labor market. Thus, the cost of additional investment in human capital is a combination of purchased inputs plus the opportunity cost of the investor's time, and the latter depends on the amount already invested in human capital and the individual's ability to effectively utilize previous investments to earn income.

The rate-of-return to this investment depends on the efficiency with which the individual can produce "earnings potential". Ben-Porath (1967) has shown that if purchased market resources are optimally combined with the investor's

---

See footnote(s) on page 143.

time, the rate-of-return on the investment must be a declining function of the amount invested in that time period.(2)

The interest rate on funds for investment (including foregone earnings) will depend on the amount that is required (the level of investment h) the amount of human capital previously accumulated (the availability of scholarships and the interest rates on loans for graduate study differ from those for undergraduate or high school study) and a broad range of other factors relating to opportunities such as economic and social background, and whether or not the investor is married.

The foregoing discussion can be illustrated diagrammatically as in Chart A.1 which illustrates the flow decisions to invest in human capital (say, schooling) by an individual in particular periods. The horizontal axis represents the flow measured in dollars (h), while the vertical axis measures the internal rate-of-return on the investment and the marginal cost of capital funds (interest rate). The d_1, d_2, and d_3 curves indicate the marginal rates-of-return to investment in periods one, two, and three respectively, where d_2 takes into account the productivity effects of investment in period one. The c_1, c_2, and c_3 curves indicate the opportunity costs of the funds to the individual in successive periods. The optimal amount of investment in, say, period two is determined by the intersection of d_2 and c_2.

In general, a two-equation simultaneous system for the ith individual can be specified as follows:

$$r_i^d = \beta_0 + \beta_1 h_i + \sum_{j=2}^{M} \beta_j X_{ij}^d + u_i^d \qquad (A.2)$$

and

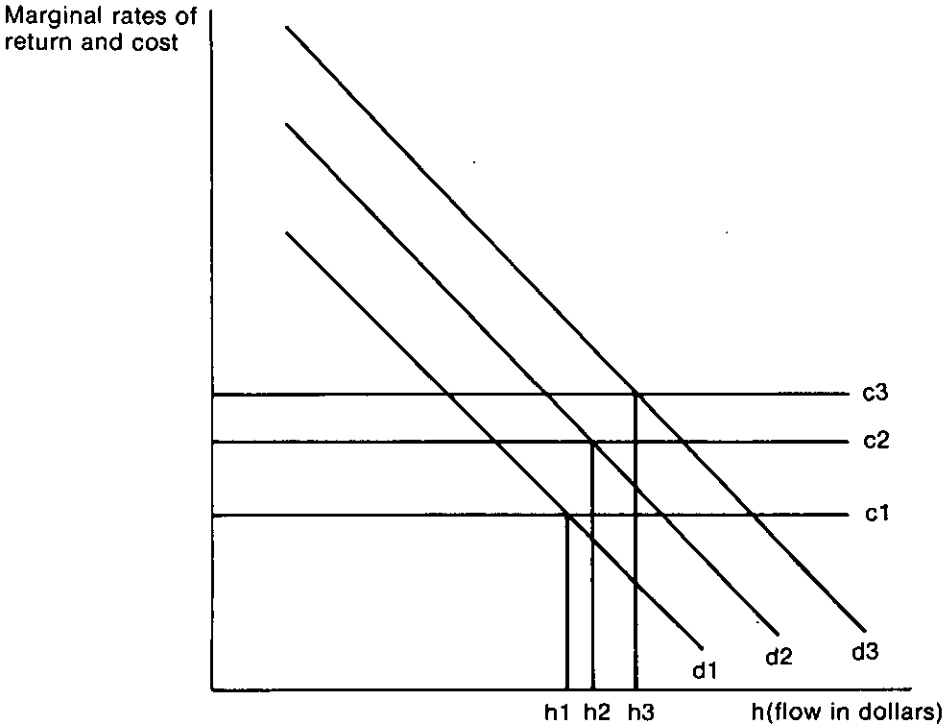
$$r_i^s = \delta_0 + \delta_1 h_i + \sum_{j=2}^{N} \delta_j X_{ij}^s + u_i^s, \qquad (A.3)$$

where the X_{ij}^s's and X_{ij}^d's represent variables affecting the position of the demand-for-investment and supply-of-funds curves respectively, u_i^s and u_i^d are statistical error terms, and the β's and δ's are unknown parameters. In equilibrium, $r_i^s = r_i^d$. These equations cannot be estimated since data on r_i and h_i are not available. How-

See footnote(s) on page 143.

Chart — A.1

Determination of Rate-of-return to Investment in Human Capital



ever, it is possible to solve for a reduced form for r_i which can be written as

$$r_i = \gamma_0 + \sum_{k=1}^{K} \gamma_k X_{ik} + u_i, \qquad (A.4)$$

where the γ's are combinations of the β's and δ's, $X_{ik}(k = 1,\ldots,K)$ contains all the X_{ij} variables in Equations A.2 and A.3 (which are assumed exogenous or predetermined) and u_i is a combination of u_i^s and u_i^d. Thus, the rate-of-return to human capital for individual i is some average (expected) rate-of-return $(\gamma_0 + \Sigma\gamma_k X_{ik})$ which depends on the levels of the X_{ik}'s plus a random component u_i. Equation A.4 is identical to Equation 3.5.

FOOTNOTES

(1) Other researchers discussing the optimal investment in human capital include Ben-Porath (1967) and Haley (1973). Rosen (1976) provides a useful summary.

(2) Ben-Porath (1967, p. 353) defined human capital as a concept analogous to "machines" in the case of physical capital. He defines a unit of human capital as the quantity of human capital required to yield some a rental rate, R, in the human capital market. Defined in this way, the stock of human capital, K, possessed by an individual is simply the individual's income potential divided by R. Ben-Porath shows that for a reasonable set of assumptions, the marginal cost of acquiring units of human capital is an increasing function of the number of units acquired per period of time. Since the rental rate per unit of K is fixed, the marginal rate-of-return on the investment must be a declining function of the amount invested in that period.

APPENDIX B

THE ESTIMATION PROCEDURE

In this appendix we discuss the derivation and estimation of Equation 3.6. Substituting Equations 3.3, 3.4 and 3.5 into 3.2 we get

$$Y_i = (\alpha_0 + \sum_j \alpha_j Z_{ij} + e_i) + [S_i + \theta\, f(EX_i)]$$

$$(\gamma_0 + \sum_k \gamma_k X_{ik} + u_i), \tag{B.1}$$

which may be rewritten as

$$Y_i = \alpha_0 + \sum_j \alpha_j Z_{ij} + [S_i + \theta\, f(EX_i)]\,[\gamma_0 + \sum_k \gamma_k X_{ik}] + \varepsilon_i, \tag{B.2}$$

where

$$\varepsilon_i = e_i + [S_i + \theta\, f(EX_i)]\, u_i = e_i + H_i u_i \tag{B.3}$$

is an error term.

Equation B.1 can be considered a nonlinear random coefficients model. It is a random coefficients model since the intercept (a_i) and the returns to human capital (r_i) are random. Both consist of a systematic component and a residual error term. It is nonlinear since both H_i and r_i depend on unknown parameters.(1) If the error terms in B.1 are collected into a common error term as in Equation B.2, the resulting error term ε_i described in Equation B.3 is heteroskedastic. If $E(e_i) = E(u_i) = E(u_i e_i) = 0$, then $E(\varepsilon_i) = 0$ and

$$\mathrm{Var}\varepsilon_i = E(\varepsilon_i^2) = \sigma_e^2 + H_i^2\,\sigma_u^2 = \sigma_i^2, \tag{B.4}$$

where H_i which is defined in (3.3) depends on the unknown parameter θ, and σ_e^2 and σ_u^2 are the constant variances of e_i and u_i, respectively, and S_i and EX_i are

---

See footnote(s) on page 148.

assumed to be nonstochastic. Thus, the model as it appears in B.2 may be viewed as a nonlinear constant coefficients model with a heteroskedastic error term. It is conceptually possible to obtain maximum likelihood estimates of the parameters of B.2 and B.4, in which case it would be possible to test the hypotheses $\alpha_j = \gamma_k = 0$ for all j and k > 0 using likelihood ratio procedures. However, maximum likelihood estimation becomes extremely nonlinear since some of the unknown parameters of B.2 also appear in B.4.

It is possible to reparameterize Equation B.2 and obtain a linear equivalent as follows:

$$Y_i = \alpha_0 + \sum_{j=1}^{J} \alpha_j Z_{ij} + \gamma_0 S_i + \lambda_0 f(EX_i)$$

$$+ \sum_{k=1}^{K} \gamma_k S_i X_{ik} + \sum_{k=1}^{K} \lambda_k f(EX_i) X_{ik} + \varepsilon_i, \qquad \text{(B.5)}$$

where the nonlinear restrictions $\lambda_k = \theta\gamma_k$ are ignored. The linear model in B.5 will reduce to the simple model in 3.1 if (as above) $\alpha_j = \gamma_k = 0$ for all j and k > 0. However, in this form it is not possible to test directly whether $\gamma_k = 0$ since combined parameters of the form $\theta\,\gamma_k = \lambda_k$ are estimated. Furthermore, ordinary least squares estimation of B.5 will be fraught with heteroskedasticity problems and will result in unbiased and consistent, but inefficient estimates of the parameters. The usual estimates of the variance-covariance matrix of the estimated parameters will be both biased and inconsistent making it difficult to test hypotheses. Thus, Equation B.5 was estimated using the following maximum likelihood procedure. Rewrite all n observations on B.5 as

$$Y = W\beta + \varepsilon, \qquad \text{(B.6)}$$

where Y and ε are n x 1 vectors with elements Y_i and ε_i, W is an n x p matrix of observations on the independent variables including a unit column, and β is a p x 1 vector of parameters to be estimated where

$$\beta' = [\alpha_0, \alpha_1, \ldots, \alpha_J, \gamma_0, \lambda_0, \ldots, \gamma_K, \lambda_K].$$

Equation B.4 can be rewritten as

$$\sigma_i^2 = (\eta + H_i^2)\sigma_u^2 = v_i \sigma_u^2 \qquad \text{(B.7)}$$

where $\eta = \sigma_e^2/\sigma_u^2$ is the unknown ratio of the variances. The assumptions above imply

$E\varepsilon = 0$. If it is further assumes that $E\varepsilon_i\varepsilon_j = 0$ for $i \neq j$ (a reasonable assumption for cross-sectional data), then

$$E\varepsilon\varepsilon' = \sigma_u^2 V, \tag{B.8}$$

where V is an n x n diagonal matrix with v_i on the main diagonal. If ε_i is assumed to be normally distributed, then the logarithm of the likelihood function can be written as

$$L = -(n/2)\ [\ln(2\pi) + \ln\sigma_u^2] - (1/2) \sum_{i=1}^{n} \ln v_i$$
$$-(1/2\sigma_u^2)\ (Y - W\beta)'V^{-1}\ (Y - W\beta). \tag{B.9}$$

It is desired to maximize L with respect to β, σ_u^2 and η. The first-order conditions can be written as follows:

$$\frac{\partial L}{\partial \beta} = -\frac{1}{\sigma_u^2}\ (W'V^{-1}W\beta\ - W'V^{-1}Y) = 0 \tag{B.10}$$

$$\frac{\partial L}{\partial \sigma_u^2} = -\frac{n}{2\sigma_u^2} + \frac{1}{2(\sigma_u^2)^2}\ (Y - W\beta)'V^{-1}\ (Y - W\beta) = 0 \tag{B.11}$$

$$\frac{\partial L}{\partial \eta} = -\frac{1}{2}\sum_{i=1}^{n}\ (\frac{1}{v_i}) + \frac{1}{2\sigma_u^2}\ (Y - W\beta)'V^{-1}V^{-1}\ (Y - W\beta) = 0 \tag{B.12}$$

Simultaneous solution of B.10 to B.12 is extremely difficult. Thus the following iterative procedure has been used. Obtain estimates of β and σ_u^2 by solving B.10 and B.11 simultaneously using an arbitrary starting value for η. Simultaneous solution of B.10 and B.11 implies

$$\hat{\beta} = (W'V^{-1}W)^{-1}W'V^{-1}Y \tag{B.13}$$

and

$$\hat{\sigma}_u^2 = \frac{1}{n}(Y'V^{-1}Y - \hat{\beta}'W'V^{-1}Y). \tag{B.14}$$

Equations B.13 and B.14 are the usual generalized least squares estimates when the V matrix is known. However, V consists of a diagonal matrix with the typical diagonal element $v_i = \eta + H_i^2$ which is not known. As mentioned above, an arbitrary starting value of $\eta = \eta_0$ can be selected. However, H_i is the level of human capital of the ith individual which is not directly observable and must be estimated. Assuming that $f(EX_i) = \theta_1 EX_I + \theta_2 EX^2$ in Equation 3.3 is a quadratic function of experience, an estimate of H_i can be obtained as

$$\hat{H}_i = [\gamma_0\ \ S_i + \gamma_0\theta_1 EX_i + \gamma_0\theta_2 EX_i]/\gamma_0 \tag{B.15}$$

and γ_0 , $\gamma_0\theta_1$, and $\gamma_0\theta_2$ are consistent OLS estimates of the parameters of equation (B.6) (i.e., Equation 3.6 of the text).

After estimates of $\hat{\beta}$ and $\hat{\sigma}_u^2$ have been obtained using Equations B.13, B.14, and B.15, these estimates can be substitued into the log likelihood function to obtain a marginal likelihood function,

$$L^* = -n/2[\ln(2\pi) - \ln \hat{\sigma}_u^2] - (1/2)\Sigma(\eta_o + \hat{H}_i^2)$$
$$- (1/2\ \sigma_u^2) \sum_{i=1}^{n} [(S_i - \sum_{k=1}^{p} \hat{\beta}_k W_{ik})Y_i/(\eta_o + \hat{H}_i^2)] . \quad \text{(B.16)}$$

The marginal likelihood function was evaluated for alternative values of η, and the value of η corresponding to the maximum of L* was used as a revised estimate. This new value, $\hat{\eta}_1$, along with revised estimates obtained from Equation B.13, were then used to obtain a new V matrix to be used in a second iteration to obtain revised estimates of β and σ_u^2. This iterative process was continued until the estimates converged to a stable value. The stopping criterion used was that the largest change in any parameter between iterations was less than 1%.

FOOTNOTE

(1) We are concerned with the generalization of the random coefficients model by Singh et al., (1976) to what they term the variable mean response model. Note that if H_i were directly observable, this would be a linear random coefficients model.

APPENDIX C

ESTIMATING FEMALE EXPERIENCE CORRECTIONS FOR CHILDBEARING

Define potential gross earnings (Y*) in year t as

$$Y_t^* = RK_t, \tag{C.1}$$

where R is the rental price per unit of human capital and K_t is the amount of human capital at time t. Actual earnings (Y) in year t are

$$Y_t = Y_t^* - C_t = Y_t^*(1 - \theta_t) = RK_t\ (1 - \theta_t), \tag{C.2}$$

where C_t is gross investment in human capital in year t measured in terms of foregone gross earnings, or $C_t = R\theta_t K_t$, where θ_t is the proportion of the existing stock of human capital devoted to the accumulation of more human capital. The change per unit time in human capital can be written as

$$\frac{dK}{dt} = \dot{K} = C_t - \delta K_t = (R\theta_t - \delta)K_t, \tag{C.3}$$

where δ is a constant rate of depreciation of the existing stock of human capital.

Now assume that a female's life beginning at age six can be divided exhaustively into three categories: i) periods of formal schooling, ii) periods of no work and no formal schooling, and iii) periods of work and no formal schooling.

During type-i periods of formal schooling, all previously accumulated human capital is devoted to the accumulation of more human capital and $\theta_t = 1$. Therefore,

$$\dot{K} = (R - \delta)K_t \text{ for } 0 \leq t \leq S, \tag{C.4}$$

where S equals the total number of years of schooling undertaken. At any point of time during formal schooling, the accumulated stock of human capital will be

$$K_t = K_o e^{(R - \delta)t} \text{ for } 0 \leq t \leq S, \tag{C.5}$$

where K_o is the initial stock of human capital.

During type-ii periods there is no investment in human capital and $\theta_t = 0$. Hence, the rate of change of human capital is

$$\dot{K} = -\delta K_t, \tag{C.6}$$

and at a given point of time during the hiatus from work and school, the human capital stock will be

$$K_t = K_1 e^{-\delta t} \quad \text{for } 0 \leq t \leq H, \tag{C.7}$$

where H represents the length of the hiatus, and K_1 is the size of the human capital stock at the beginning of the hiatus.

Finally, during type-iii periods, the existing stock of human capital can be allocated to the accumulation of more human capital through on-the-job training, to earning income through market activities, or to some combination of the two. Thus, during this period $0 \leq \theta_t \leq 1$. Considerations from the theory of the optimal accumulation of human capital suggest that θ will decline over the working life of the individual (Haley, 1973) and it is usually assumed that investment in human capital ceases toward the end of the working life (i.e., θ becomes zero). Following Mincer (1974), we impose the restriction that θ declines linearly over the working life of the female and

$$\theta_t = a + bt \quad \text{for } 0 \leq t \leq T, \tag{C.8}$$

where T is the total length of working life in years, and a and b are parameters such that $0 \leq a \leq 1$ and $b < 0$. Thus, during the working life, the rate of change of human capital is

$$\dot{K} = (\theta_t R - \delta)K_t = [R(a + bt) - \delta]K_t. \tag{C.9}$$

The corresponding stock of human capital at any point in time will be

$$K_t = K_2 \exp\{(Ra - \delta)t + Rbt^2/2\}, \tag{C.10}$$

where K_2 is the stock of human capital at the beginning of the work period.

Given the above assumptions about the rental value, depreciation and accumulation of human capital, it is easy to show that the stock of human capital at a given point of time is independent of the order of the schooling-hiatus-experience sequence which preceded it, and that the stock depends only on the length of time involved in each phase. Human capital accumulation during the working life depends on the stock of human capital at the beginning of the working period, actual elapsed working time, and the constant parameters R, a, b, and δ. Given these assumptions, the accumulated capital stock is independent of the sequence regardless of whether there were interruptions in schooling or work experience for childbearing or for any other reason. Therefore, for simplicity of exposition but without loss of generality, we assume that the sequence is schooling-hiatus-continuous experience.

Consider a woman who has S years of schooling, a hiatus of H years, and X years of continuous experience. Substituting Equations C.5 and C.7 into Equation C.10, it can be shown that the stock of human capital is

$$K_X = K_o \exp\{(R - \delta)S - \delta H + (Ra - \delta)X + RbX^2/2\}. \tag{C.11}$$

Substituting Equations C.11 and C.8 into C.2, net earnings are

$$Y = \{R(1 - a - bX)K_o\}\exp\{(R - \delta)S - \delta H + (Ra - \delta)X + RbX^2/2\}. \tag{C.12}$$

Taking natural logarithms of both sides of C.12 results in

$$\ell nY = \ell n(RK_o) + \ell n\,(1 - a - bX) + (R - \delta)S - \delta H + (Ra - \delta)X + (Rb/2)X^2 \tag{C.13}$$

Since a + bX is small, using a Taylor series expansion it can be shown that $\ell n(1 - a - bX) \simeq - a - bX$. Hence Equation C.13 can be approximated as

$$E = \ell nY \simeq \beta_0 + \beta_1 S - \delta H + \beta_2 X + \beta_3 X^2 \tag{C.14}$$

where $\beta_0 = \ell n\, RK_o - a$, $\beta_1 = (R - \delta)$, $\beta_2 = (Ra - \delta - b)$, and $\beta_3 = Rb/2$.

Unfortunately, Equation C.14 cannot be estimated directly for females since the census data do not provide observations on H and X. However, it is possible to obtain indirect estimates of H and X under the assumption that H is a function of the number of children ever born (N). Let such a function be approximated by a quadratic as

$$H = \alpha_1 N + \alpha_2 N^2, \tag{C.15}$$

where $\alpha_1 > 0$ and $\alpha_2 \gtrless 0$. Define potential experience X*, for females as age (A) minus years of schooling minus six. If the hiatus for childbearing and rearing accounts for the entire difference between actual labor force experience and potential experience, then

$$X^* = X - H = A - S - 6, \qquad (C.16)$$

Substituting Equation C.15 for H and X* for X into C.14, the following estimable equation is obtained:

$$E = \beta_0^* + \beta_1^* S + \Psi_1 N + \Psi_2 N^2 + \beta_2^* X^* + \beta_3^* X^{*2}, \qquad (C.17)$$

where $\Psi_1 = -\delta\alpha_1$ and $\Psi_2 = -\delta\alpha_2$. From C.14 and C.15, it is evident that if $\hat{\Psi}_1$ and $\hat{\Psi}_2$ are estimates of the parameters of C.17, then

$$\hat{\Psi}_1 N + \hat{\Psi}_2 N^2 = -\hat{\delta H} \qquad (C.18)$$

is an estimate of the decline in the natural logarithm of potential earnings due to depreciation of the human capital stock and foregone investment opportunities during the childbearing hiatus. Define Z as the additional amount of potential work experience required to offset the loss in earning power due to the birth of a given number of children. Given the estimated coefficients of potential experience and potential experience squared in C.17 and using C.18 we obtain

$$\hat{\beta}_2^* Z + \hat{\beta}_3^* Z^2 = -\hat{\Psi}_1 N - \hat{\Psi}_2 N^2. \qquad (C.19)$$

Equation C.19 can be solved for the required additional experience to offset the loss in earnings as a function of the number of children as[(1)]

$$Z = \frac{-\hat{\beta}_2^* + \sqrt{\hat{\beta}_2^{*2} + 4\hat{\beta}_3^*(\hat{\Psi}_1 N + \hat{\Psi}_2 N^2)}}{2\hat{\beta}_3^*}. \qquad (C.20)$$

However, Z is an estimate of the length of the hiatus in work experience, H, and an estimate of actual experience, X, can be obtained using C.20 and C.16. Hence, Equation C.14 can be estimated by replacing H and X by the estimates $\hat{H} = Z$, and $\hat{X} = X^* - \hat{H}$. This involves first estimating Equation C.17, obtaining estimates of H via Equation C.20, and then estimating Equation C.14 at the second stage.[(2)]

See footnote(s) on page 153.

FOOTNOTES

(1) Equation C.14 has two solutions corresponding to the positive and negative square roots. The positive root was used since using the negative square root resulted in negative estimates for the hiatus for every ethnic group.

(2) While this procedure is probably an improvement over not making any correction for childbearing or just using first stage results, it should be noted that the first stage of this procedure involves an errors in variables problem since potential experience measures actual experience with an upward bias (i.e., $X^* \geq X$). Thus, the first stage estimates will not be consistent. As a consequence of this, the second stage estimates will not be consistent either.

APPENDIX D

DEFINITION OF VARIABLES USED

D.1. Binary Variables

(1 if the characteristics is listed, 0 otherwise)

Ethnic Groups

ETH1 — French

ETH2 — Negro or West Indian

ETH3 — Austrian, Finnish, German, Dutch or Scandinavian

ETH4 — Czech, Hungarian, Polish Russian, Slovak, or Ukranian

ETH5 — Italian

ETH6 — Jewish

ETH7 — Chinese or Japanese

ETH8 — Native Indian

ETH9 — Other and Unknown

(British Isles is the excluded category.)

Regions

REG1 — Maritime Provinces

REG2 — Quebec

REG3 — Ontario

REG4 — Praire Provinces

(British Columbia is the excluded region.)

Community Type

RES1 — urban area of 30,000 persons or more

RES2 — rural non-farm

RES3 — rural farm

(Urban area of less than 30,000 persons is the excluded category.)

Period of Immigration

PIM1 — before 1946

PIM2 — 1946-60

PIM3 — 1961-70

(Canadian-born is the excluded category.)

Occupations

OC1 — managerial, administrative, and related

OC2 — natural science, engineering and mathematics

OC3 — social science and related

OC4 — religion

OC5 — teaching and related

OC6 — medicine and health

OC7 — artistic, literary, recreation, and related

OC8 — other[1]

OC9 — sales

OC10 — service

OC11 — farming, horticulture, and animal husbandry

OC12 — other primary occupations (fishing, hunting, trapping, forestry, mining)

OC13 — processing

OC14 — machining, product fabrication, assembling, and repairing

OC15 — construction trade occupations

OC16 — transport equipment operating

(1) Materials handling and related occupations, other crafts and equipment operating occupations (printing and utilities) and occupations not elsewhere classified.

(Clerical occupations is the excluded category.)

Industry

IN1 — agriculture

IN2 — forestry

IN3 - fishing and trapping

IN4 - mines, milling, quarries, and oil fields

IN5 - construction

IN6 - transportation, communication and other utilities

IN7 - trade

IN8 - finance, insurance, and real estate

IN9 - commumity, business and personal services

IN10 - public administration and defense

(Manufacturing is the excluded caterogy.)

Education and Training

SQD – highest level of elementary or secondary school attended was in the Yukon, the Northwest Territories, or outside Canada

COL – completed a vocational training course of three months or longer duration

V – completed an apprenticeship program of six months or longer duration

HS – highest level of schooling attended was Grade 9 through 13

UN – attended university

DEG – received a university degree

ASCH– attended school full or part-time after September 1970

Working Time Per Week

PTHRS – usually worked less than 35 hours per week

TM1 – usually worked part-time weeks in 1970

OTHRS – usually worked 45 or more hours per week

OTSE – self-employed and usually worked 45 or more hours per week

Miscellaneous

HH — head of census family

MS1 — currently married

MS2 — currently widowed, divorced or separated

EM — ever-married

SE — self-employed

IMIG — living in different province or county within Canada in 1966

MIS — married prior to completing school, i.e., if age at first marriage is less than or equal to YSCH + 6

MIG — living outside Canada in 1966

LNO2 — ability to speak both English and French

D.2. Continuous Variables

AGE — age at last birthday

YSCH — years of schooling attended (including full-time vocational courses)

EX — years of potential labor force experience (EX = (AGE — YSCH — 6)) if negative set equal to zero

EE — years of potential experience squared

E — natural logarithm of earned income: wages and salaries plus self-employment income

LWW — the natural logarithm of weeks worked which was constructed as follows:

If weeks worked during 1970 was --

1 — 13 weeks, LWW = $\ln 7$

14 — 26 weeks, LWW = $\ln 20$

27 — 39 weeks, LWW = $\ln 33$

40 — 48 weeks, LWW = $\ln 44$

49 — 52 weeks, LWW = $\ln 51$

LWR – the natural logarithm of the implied weekly wage rate which was defined as E – LWW

LWS – the natural logarithm of wages and salaries

WAGE – 1970 mean earnings in the individual's occupation and province

KLR – 1970 capital-to-labor ratio in the individual's industry and province

SQ – school quality. Measured by operating-expenditures per-student-enrolled in constant 1971 dollars for the year and province in which the individual attended highest level of elementary or secondary school

D.3. Continuous Variables Defined for Females Only

NCB – number of children ever born

NC2 – number of children ever born squared

HAT – predicted hiatus in work experience due to birth of children

EXS – years of potential experience adjusted for hiatus due to pregnancies and birth of children

EES – years of potential experience adjusted for hiatus squared

APPENDIX E

INTERPRETATION OF THE COEFFICIENTS OF THE FULL MODEL

This Appendix discusses in detail the coefficients estimated for the full annual earnings model which are presented in Table E.4. The discussion focuses on the net (or partial) effect of each variable on the employment income received by individuals and compares the anticipated direction of these effects with the signs of the estimated coefficients. We begin by discussing the coefficients for the human capital variables and the interaction of these variables with the X's of Chapter 3; i.e., the factors affecting the rate-of-return to human capital. This is followed by interpretation of the coefficients for the personal attributes and the market effects variables. These are the Z's discussed in Chapter 3. Since many of these same variables are proxies for factors affecting the rate-of-return to human capital as well as proxies for factors affecting earnings directly, these variables serve as X's as well.

E.1. Human Capital Effects and Interactions

E.1.1. Years of Schooling (YSCH)

The partial effect of each additional year of schooling on earnings is expected to be positive for each ethnic group. This partial effect, or the percentage rate-of-return to schooling, is calculated as

$$100\left(\frac{\partial E_i}{\partial YSCH_i}\right) = 100\,(a_{1k} + a_2 SQ_i + a_3 KLR_i + a_4 V_i + a_5 MIS_i), \qquad (E.1)$$

where a_{1k} is the estimate of the schooling coefficient for the kth ethnic group (the appropriately adjusted British coefficient) and a_j (j=2,...,5) are respectively the coefficients of the interaction of years of schooling with schooling quality (SQ), the provincial capital/labor ratio of the industry of employment (KLR), a dummy variable which is 1 if the individual participated in an apprenticeship program (V), and a dummy variable which is one if the individual was married in school (MIS).(1) The estimated coefficient of years of schooling (a_k) for the British (reference) group is 0.0571, and for the French group (ETH1), the coefficient is

---

See footnote(s) on page 181.

TABLE E.1. Effects on the Location of the Earnings/Experience Peak of Various Factors[1]

Ethnic group	Married in school	Self-employed	Apprenticeship training	Attended high school	Attended university	University degree	Capital/labor ratio
French	earlier	earlier	earlier	earlier	earlier	later	later
North European	earlier	earlier	earlier	earlier	earlier	later	later
East European	later	earlier	earlier	earlier	earlier	later	later
Italian	later	earlier	earlier	earlier	earlier	later	later
Jewish	earlier	earlier	earlier	earlier	earlier	later	later
Oriental	later	earlier	earlier	earlier	earlier	later	earlier
Native Indian	later	earlier	later	later	later	later	earlier
Other Unknown	earlier	earlier	earlier	earlier	earlier	later	later
British	earlier	earlier	earlier	earlier	earlier	later	later

(1) Evaluated using Equation E.5 and assuming that the capital/labor ratio is $23.07 per worker. This is the arithmetic average of the KLR's for the sample used to estimate the full model.

Source: Computed from Table E.4 and Public Use Sample Tape.

TABLE E.2. Maximum Likelihood Estimates of the Reduced Model with the Logarithm of Weekly Wage Rate (LWR) as the Dependent Variable (1)

Personal attributes

Intercept	ETH1	ETH2	ETH3	ETH4	ETH5	ETH6	ETH7	ETH8	ETH9	MSI	MIS	LNO2
3.887	.3438	-1.769	.270	.0383	.500	-.225	-.0633	.811	.237	.0526	-.866	.0411
(.1017)	(.0778)	(.666)	(.109)	(.136)	(.186)	(.265)	(.361)	(.416)	(.161)	(.282)	(.0863)	(.0200)

OTHRS	OTSE	PTHRS	PIM1	PIM2	PIM3	MIG	IMIG	SE
-.0672	-.2129	-.499	.0219	.0576	.00664	-.0197	.0136	.158
(.0170)	(.0463)	(.0262)	(.0398)	(.0347)	(.0528)	(.0447)	(.0197)	(.0838)

Direct human capital effects

COL	V	DEG	SQ	SQD
-.0799	.1753	.0580	-.000541	-.00557
(.0277)	(.1327)	(.0599)	(.000333)	(.0803)

Human capital interactions

	BRITISH (2)	ETH1	ETH2	ETH3	ETH4	ETH5	ETH6	ETH7	ETH8	ETH9	MIS	SE
YSCH	.0676 (.00770)	-.0204 (.00552)	.1159 (.0384)	-.0183 (.00758)	-.00975 (.00933)	-.0250 (.0124)	-.00356 (.0172)	.00363 (.0226)	-.0680 (.0302)	-.0278 (.0102)	.0305 (.00540)	--
EX	.0337 (.00351)	-.0129 (.00367)	.0152 (.0410)	-.00505 (.00474)	.00876 (.00597)	-.00816 (.00900)	.0415 (.0109)	.00366 (.0140)	-.0104 (.0209)	-.000179 (.00775)	.0297 (.00415)	-.0119 (.00556)
EE	-.000480 (.0000583)	.000218 (.0000688)	-.000133 (.00103)	.0000324 (.0000191)	-.000235 (.000114)	.0000183 (.000167)	-.000640 (.000215)	.00000951 (.000240)	-.000103 (.000368)	.0000119 (.000142)	-.000486 (.0000842)	.0000560 (.0000860)

	V	HS	UN	DEG	SQ	SQD
YSCH	-.0141 (.00944)	--	--	--	.0000326 (.0000283)	-.00537 (.00713)
EX	.00643 (.00651)	.00606 (.00238)	.00575 (.00524)	.00382 (.00773)	--	--
EE	-.000158 (.000126)	-.000225 (.0000547)	-.000299 (.000136)	.000156 (.000196)	--	--

-- not applicable.

(1) Standard errors in parentheses.

(2) British is the reference group and all interactions should be interpreted as deviations from this group.

Source: Public Use Sample Tape.

TABLE E.3. Maximum Likelihood Estimates of the Reduced Model with the Logarithm of Annual Earnings (E) as the Dependent Variable[1]

Personal attributes

Intercept	ETH1	ETH2	ETH3	ETH4	ETH5	ETH6	ETH7	ETH8	ETH9	MSI	MIS	LNO2
4.413	.365	-01.839	.259	.0873	.571	-.183	.0307	.693	.237	.0571	-1.044	.0426
(.1097)	(.0782)	(.676)	(.110)	(.137)	(.185)	(.271)	(.365)	(.403)	(.160)	(.0286)	(.0873)	(.0200)

OTHRS	OTSE	PTHRS	PIM1	PIM2	PIM3	MIG	IMIG	SE	LWW
-.0707	-.187	-.545	.0194	.0628	.000979	-.0122	.0193	.1685	.847
(.0170)	(.0466)	(.0260)	(.0401)	(.0346)	(.0529)	(.0452)	(.0198)	(.0831)	(.0141)

Direct human capital effects

COL	V	DEG	SQ	SQD
-.0773	.2084	.0785	-.000479	.0107
(.0284)	(.1336)	(.0610)	(.000323)	(.0785)

Human capital interactions

	BRITISH[2]	ETH1	ETH2	ETH3	ETH4	ETH5	ETH6	ETH7	ETH8	ETH9	MIS	SE
YSCH	.0703	-.0230	.120	-.0183	-.0126	-.0284	-.00856	-.00184	-.0609	-.0282	.0374	--
	(.00763)	(.00567)	(.0394)	(.00779)	(.00960)	(.0126)	(.0179)	(.0233)	(.0297)	(.0193)	(.00554)	
EX	.0352	-.0127	.0142	-.00359	.00783	-.0113	.0421	.00260	-.00908	-.000330	.0392	-.0158
	(.00346)	(.00360)	(.0415)	(.00467)	(.00587)	(.00883)	(.0198)	(.0138)	(.0201)	(.00761)	(.00417)	(.00547)
EE	-.000500	.000205	-.0000480	.00000286	-.000233	.0000617	-.000642	-.00000458	-.000112	.0000190	-.000620	.000132
	(.0000569)	(.0000670)	(.00104)	(.0000897)	(.000111)	(.000162)	(.000210)	(.000230)	(.000351)	(.000137)	(.0000827)	(.0000865)

	V	HS	UN	DEG	SQ	SQD
YSCH	-.0140	--	--	--	.0000244	-.00775
	(.00964)				(.0000284)	(.00171)
EX	.00383	.00802	.00803	.000884	--	--
	(.00647)	(.00236)	(.00543)	(.00821)		
EE	-.000122	-.000274	-.000352	.000213	--	--
	(.000125)	(.0000544)	(.000143)	(.000210)		

-- not applicable.

(1) Estimated standard errors in parentheses.

(2) British is the reference group and all interactions should be considered as deviations from this group.

Source: Public Use Sample Tape.

TABLE E.4. Maximum Likelihood Estimates of the Full Model with the Logarithm of Annual Earnings (E) as the Dependent Variable(1)

Personal attributes and market effects

Intercept	ETH1	ETH2	ETH3	ETH4	ETH5	ETH6	ETH7	ETH8	ETH9	MAR	MSI	LNO2	SE
4.280	.285	-2.027	.235	.0558	.320	-.313	.104	.662	.0719	.0430	-.949	.0323	.210
(.162)	(.0774)	(.00658)	(.107)	(.113)	(.179)	(.264)	(.354)	(.389)	(.156)	(.0276)	(.0855)	(.0195)	(.0817)

OTHRS	OTSE	PTHRS	PIM1	PIM2	PIM3	MIG	IMIG	OC1	OC2	OC3	OC4	OC 5	OC6
-.00362	-.0554	-.479	-.0118	.042	-.0214	-.0421	.00774	-.454	-.123	-.306	-.484	-.0885	-.519
(.0170)	(.0457)	(.0256)	(.0383)	(.0036)	(.0515)	(.0441)	(.0194)	(.130)	(.0644)	(.105)	(.122)	(.0726)	(.151)

OC7	OC8	OC9	OC10	OC11	OC12	OC13	OC14	OC15	OC16	IN1	IN2	IN3	IN4
-.0327	.103614	-.0557	.00541	-.0153	.0777	.0693	.0133	.0182	.00899	-.195	.132	-.171	.0788
(.0660)	(.0317)	(.0382)	(.0342)	(.0768)	(.0552)	(.0372)	(.0311)	(.0357)	(.0355)	(.0731)	(.0618)	(.102)	(.0652)

IN5	IN6	IN7	IN8	IN9	IN10	REG1	REG2	REG3	REG4	RES1	RES2	RES3	WAGE
-.0813	-.0353	-.108	-.0195	-.127	-.0131	-.163	-.137	-.0817	-.105	.0577	-.0489	-.0160	.000109
(.0331)	(.0456)	(.0242)	(.0430)	(.0271)	(.0383)	(.0466)	(.0357)	(.0249)	(.0308)	(.0181)	(.022578)	(.0372)	(.000016)

KLR	LWW
.00417	.827
(.00156)	(.0138)

Direct human capital effects

COL	V	DEG	SQ	SQD
-.0837	.176	.0166	-.00137	-.133
(.0276)	(.129)	(.0608)	(.000332)	(.0769)

Human capital interactions

	BRITISH (2)	ETH1	ETH2	ETH3	ETH4	ETH5	ETH6	ETH7	ETH8	ETH9	MIS	SE
YSCH	.0571	-.0155	.139	-.0138	-.00990	-.0142	.00166	.0100	-.0475	-.0179	.0350	--
	(.00777)	(.00553)	(.0383)	(.00756)	(.00933)	(.0122)	(.0174)	(.0226)	(.0286)	(.0101)	(.00541)	
EX	.0270	-.0111	.00483	-.00422	.00961	-.00774	.0353	.00592	-.0179	.00169	.0341	-.0172
	(.00377)	(.00347)	(.0403)	(.00450)	(.00564)	(.00848)	(.0104)	(.0133)	(.0194)	(.00733)	(.00406)	(.00530)
EE	-.000425	.000184	.000347	.000030	-.000262	.000051	-.000508	-.000093	.000001	.000026	-.000526	-.000194
	(.000062)	(.000064)	(.00101)	(.000086)	(.000106)	(.000155)	(.000202)	(.000221)	(.000336)	(.000131)	(.000079)	(.000084)

	V	HS	UN	DEG	SQD	SQ	KLR
YSCH	-.0124	--	--	--	-.0101	.000019	-.000099
	(.00933)				(.00694)	(.000027)	(.000094)
EX	.00354	.00648	.00555	.00464	--	--	-.000152
	(.00623)	(.00227)	(.00528)	(.00802)			(.000063)
EE	-.000128	-.000230	-.000296	.000150			.000003
	(.000120)	(.000052)	(.000139)	(.000204)			(.000001)

-- not applicable.

(1) Standard errors appear below estimated coefficients in parentheses.

(2) British is the reference group. All interactions should be interpreted as deviations from this group.

Source: Public Use Sample Tape.

0.0571 - 0.0155 = 0.0416. Thus the rate-of-return to schooling for the French group is lower than for the British, ceteris paribus. The actual rate-of-return to years of schooling must be modified for the level of schooling quality, the capital/labor ratio of the industry of employment, and whether or not the individual married in school or participated in an apprenticeship program.

The coefficient of the YSCH·SQ interaction is .000019, which implies that as schooling quality (measured as expenditure per enrolled pupil) increases, the rate-of-return to education also increases although this effect is not statistically significant. The coefficient of the YSCH·KLR interaction is -0.000099, which implies that the rate-of-return to education is lower in relatively more capital intensive industries. The coefficient of the YSCH·V interaction is -0.0124, this implies that those individuals who participated in an apprenticeship program earn a lower rate-of-return on their investment in formal education than those who did not. The coefficient of the YSCH·MIS interaction is 0.0350 and highly significant. This implies that for those individuals who married in school, the return to the last unit of investment is substantially higher than for those who were not married in school. This is consistent with the argument, presented in Chapter 6, that individuals who married in school would have a higher opportunity cost of funds for investment in formal education, and hence would require a higher rate-of-return to this investment in order to justify its greater cost.

The rates-of-return to education by ethnic group are compared in Chapter 5.

E.1.2. Experience (EX) and Experience Squared (EE)

Human capital theory suggests that post-school investments in on-the-job training will initially be positive, will decline during the working life of the individual, and will eventually become zero. This suggests that earnings will increase at a decreasing rate, and may actually decline in the latter part of the working life if net investment in human capital is negative because the depreciation of human capital exceeds gross investment. This implies that the coefficient for experience should be positive and the coefficient for experience squared should be negative. This was the case for every ethnic group.

The percentage rate-of-return to experience for the k^{th} ethnic group can be calculated as

$$100 \left(\frac{\partial E_i}{\partial EX_i}\right) = 100 \, (b_{1k} + 2c_{1k}EX + b_2MIS + 2c_2MIS \cdot EX$$

$$+ b_3SE + 2c_3SE \cdot EX + b_4KLR + 2c_4KLR \cdot EX$$

$$+ b_5V + 2c_5V \cdot EX + b_6HS + 2c_6HS \cdot EX$$

$$+ b_7UN + 2c_7UN \cdot EX + b_8DEG + 2c_8DEG \cdot EX), \qquad (E.2)$$

where b_{1k} and c_{1k} are the coefficients of experience and experience squared for the k^{th} group (the appropriately adjusted British coefficient). The $b_j (j=2,...,8)$ are coefficients for interaction terms between experience and married in school (MIS), self-employment (SE), KLR, V, a dummy variable which is 1 if the individual attended high school (HS), a dummy variable which is 1 if the individual attended university (UN), and a dummy variable which is 1 if the individual obtained a university degree (DEG). The $c_j (j=2,...,8)$ are coefficients for interaction terms between experience squared and these same variables. Equation E.2 can be rewritten as

$$100 \left(\frac{\partial E}{\partial EX}\right) = 100 \left(\sum_{j=1}^{8} b_j + 2EX \sum_{j=1}^{8} c_j X_j\right), \qquad (E.3)$$

where the X_j are the variables interacting with experience and experience squared. Earnings will peak when $\partial E / \partial EX = 0$, which has the solution

$$EX^* = \Sigma b_j / 2\Sigma c_j X_j = - B/2C, \qquad (E.4)$$

where $B > 0$ and $C < 0$. Since $\partial EX^*/\partial B = -1/2C > 0$, and $\partial EX^*/\partial C = B/2C^2 > 0$, any increase in the coefficients of either experience or experience squared tends to delay the peak in the earnings-experience profile. The effect on the location of peak earnings in the experience-earnings profiles of a simultaneous change in the coefficient of both experience and experience squared can be determined from the sign of

$$dEX^* = \frac{B}{2C} \left(\frac{dC}{C} - \frac{dB}{B}\right), \qquad (E.5)$$

which will be positive if both dB and dC are positive, and will be negative if both dB and dC are negative. If dB and dC are of opposite sign, the peak earnings will be later if $\left(\frac{dC}{C} - \frac{dB}{B}\right) < 0$. The sign of this expression depends on the ethnic group, the capital intensity of the industry of employment, whether the individual was married in school, etc. Some of the results are summarized in Table E.1.

For the British reference group, $b_1 = 0.0270$ and $c_1 = -0.000425$. Ignoring all interaction terms, the earnings profile would increase for the first 32 years of experience ($\partial E/\partial EX = 0.027 - 0.00085EX^* = 0$ implies $EX^* = 31.8$) and decrease thereafter.(2) For a male with eight years of schooling, this, in turn, implies that the age-earnings profile will peak at age 46 (= 32 + 8 + 6).

Insofar as formal schooling teaches learning skills, better educated individuals will probably learn more quickly on the job (increasing B) and, consequently, are likely to accumulate more post-school human capital. But insofar as more formal education increases earning capacity, it increases the opportunity cost of on-the-job training; and insofar as on-the-job training takes time, it reduces the potential working life over which returns to on-the-job training can be received. Together these two effects will tend to reduce the time extensiveness of investment in on-the-job training and imply an earlier cessation of on-the-job training investments (decreasing C). Thus, the earnings of relatively more educated individuals are expected to increase more rapidly with experience, with the peak in earnings coming earlier in terms of experience. From this argument we expect the coefficients for the interaction of experience and schooling level (EX·HS and EX·UN) to be positive, and the coefficients for the interaction of experience squared and schooling levels (EE·HS and EE·UN) to be negative. All four of these coefficients turn out to have the expected signs. From Table E.1 it is obvious that for individuals who attended high school or university, the earnings peak is reached with less experience than for individuals who never got beyond Grade 9, with the exception of Native Indians whose earnings peak later for individuals who attended high school or university.

The dummy variable for obtaining a degree (DEG) can be considered a proxy for ability and may also serve as a screening variable. The coefficients for the interaction of DEG with experience and experience squared are both positive. This implies that the earnings profile increases more rapidly and peaks later for degree holders than for non-degree holders with equivalent years of education.

An apprenticeship program can be viewed as a particularly intensive form of on-the-job training which increases the return to experience and shortens the

---

See footnote(s) on page 181.

years over which there is investment in on-the-job training. Thus, we would expect the coefficient for its interaction with years of experience (V·EX) to be positive, and its inter-action with experience squared (V·EE) to be negative. These coefficients turn out to have the anticipated sign, and apprenticeship training results in an earlier peak in the earnings-experience profile for all ethnic groups except Native Indians.

Individuals who get married while attending school can be viewed as having a high opportunity cost for funds to spend on human capital investments. As such, we would expect them to only make human capital investments that have a relatively high rate-of-return. Therefore, the return they get from any schooling should be higher than that of individuals who did not marry in school, and consequently, we expect the coefficient for the interaction of schooling and married-in-school to be positive. This was discussed above. Similarly, these individuals would be expected to be more diligent in learning skills on the job, investing more intensively in each year, but less extensively over time. This implies that the inter-action of years of experience and married-in-school (EX·MIS) should have a positive coefficient, and the interaction of years of experience squared and married-in-school (EE·MIS) should have a negative coefficient. These coefficients turn out to have their expected signs and are both significant at the 1% level. From Table E.1, it is apparent that for individuals who married in school, earnings tend to peak at a lower level of experience for the French, North European, Jewish, Other and Unknown, and British groups, but tend to peak later in life for the East European, Italian, Oriental, and Native Indian groups.

The coefficients for the interaction of the dummy variable for self-employment (SE) with both experience and experience squared are negative. This implies that earnings for self-employed individuals rise more slowly and peak earlier than for individuals working for wages or salaries.

For individuals employed in relatively more capital intensive industries, earnings tend to increase less rapidly with experience since the coefficient of EX·KLR is negative. However, the coefficient of the interaction between experience squared and KLR is positive tending to increase the length of time over which earnings increase. In Table E.1 it can be seen that for individuals employed in relatively more capital intensive industries, earnings will peak later in life for all ethnic groups except Orientals and Native Indians.

E.1.3. Direct Human Capital Effects

Some of the variables discussed above that interacted with years of schooling and experience have effects over and above these interaction effects. We now examine the total effects of these variables. In the case of a dummy variable (say D), the total effect can be evaluated as

$$\Delta E = E_{D=1} - E_{D=0} = d_2 YSCH + d_3 EX + d_4 EE + d_1, \qquad (E.6)$$

where D is assumed to have entered directly with coefficient d_1, and interacted with YSCH, EX, and EE with estimated coefficients d_2, d_3, and d_4 respectively. All other variables except D have been held constant. Since E is the natural logarithm of annual earnings, the change in E is approximately equal to the proportionate change in annual earnings.

Using E.6 we get the following proportionate changes in annual earnings:

$$\Delta E_{COL} = -0.0837$$

$$\Delta E_{V} = 0.176 - 0.0124YSCH + 0.0035EX - 0.000128EE$$

$$\Delta E_{DEG} = 0.0166 + 0.0046EX + 0.000150EE$$

$$\Delta E_{MIS} = -0.949 + 0.0350YSCH + 0.0341EX - 0.000536EE$$

$$\Delta E_{SE} = 0.210 - 0.0172EX - 0.000194EE$$

For the continuous variables, schooling quality and the capital/labor ratio, the analogous expressions are the following:

$$\partial E/\partial SQ = -0.00137 + 0.000019YSCH$$

$$\partial E/\partial KLR = 0.00417 - 0.000099YSCH - 0.000152EX + 0.00003EE$$

These expressions have the following interpretation. Those individuals who had part of their formal education as vocational training (COL=1) receive approximately 8.4% less income at all levels of experience and education than do their counterparts with equivalent years in an academic program (COL=0). Individuals who participated in an apprenticeship program (V=1) receive 17.6% more income as a consequence of their training, but this is offset, in part, by a lower return to YSCH of approximately 1.2% less income per year of schooling. Thus, an individual just entering the labor force with apprenticeship training and having ten years of schooling would receive 17.6 - 10 (1.24) = 5.2% more than an individual without apprenticeship training. Apprenticeship training also has an effect on the return to experience with earnings rising more rapidly initially and peaking earlier for all ethnic groups except Native Indians. University degree holders receive 1.7% more income upon starting work than non-degree holders with comparable years of education, and have higher returns to experience than non-degree holders at all levels of experience. Individuals married in school tended to receive 95% less income on entering the labor force, but this is offset in part by a 3.5% increase in earnings for each year of formal education. Thus, someone who married in school with ten years of schooling would receive a starting salary which was 95 - 3.5 (10) = 60% less than an individual entering the labor force with 10 years of schooling but who did not marry in school. As mentioned above, individuals married in school also invest more intensively in on-the-job training with earnings peaking earlier for most ethnic groups. Self-employed individuals tend to earn 21% more than individuals working for wages and salaries with comparable qualifications, but self-employed individuals have a flatter earnings-experience profile and earnings peak at a lower experience level for all ethnic groups.

An increase in school quality (SQ), measured as expenditure per pupil, tends to increase the return to education but there is also a negative offsetting direct effect which swamps the effect of the increased return to education. The sign of the coefficient of SQ had been expected to be positive. Finally, individuals employed in more capital intensive industries tend to earn higher incomes, but this is offset by lower returns to education and flatter earnings-experience profiles which peak later for all ethnic groups except the Oriental and Native Indians.

E.2. Personal Attributes and Market Effects

E.2.1. Ethnic Groups

Individuals from different ethnic or cultural backgrounds may have very different views of success or achievement and the acquisition of material goods, different attitudes toward work activity, on-the-job training, formal schooling and the selection of courses of study and occupations. Furthermore, it may also be the case that members of certain groups face various forms of discrimination, some of which may take the form of less pay for the same work, or of entry barriers to trade unions, to apprenticeship programs or to jobs which have a high skill-learning content. The members of certain groups in certain locations may find, or have found it more difficult than others to obtain financing for their education, or even to gain admission to certain educational programs (particularly professional schools). Phenomena of this sort will <u>ceterbis paribus</u> tend to shift the intercept term for the affected ethnic groups. This will result in coefficients for the ethnic dummies which are significantly different from zero, producing intercepts different from the British Isles reference group. Insofar as these effects operate through the cost of education, entry into particular programs, or even diligence in school, they will also result in coefficients for ethnic-years-of-schooling ($YSCH \cdot ETH_i$) interactions which are significantly different from zero. Where these effects operate through union entry, admission to apprenticeship programs or jobs with various skill learning contents, they will shift earnings-experience profiles, producing significant coefficients for the experience-ethnic interaction terms ($EX \cdot ETH_i$) and ($EE \cdot ETH_i$). The total effect of these ethnic variables is discussed in Chapter 5.

As concerns the direct effects of ethnicity or race alone, after controlling for their interactions with human capital variables, other personal attributes and all market effects variables, we find only positive deviations from the British reference group. Being in the French Group (ETH1) implies 28% higher earnings than the British Group; being in the North European Group (ETH3) implies 25% higher earnings; being in the Italian Group (ETH5) implies 33% higher earnings; and (most surprisingly) being in the Native Indian Group (ETH8) implies 66% higher earnings.(3) While these results may appear paradoxical at first, it should be remembered that among other things, these coefficients refer to the earnings of individuals age 15 and over, with zero years of schooling and experience. The East European (ETH4), Jewish (ETH6), and the other and unknown groups (ETH9) all had direct partial effect coefficients which were not significantly different from zero at the 20% level, implying that for zero years of schooling and experience, these groups have earnings that are not significantly different from the British Isles Group.

See footnote(s) on page 181.

When we consider the interaction effects of years of schooling and ethnicity, four groups (excluding the Negro/West Indian Group) seem to have significantly different returns to formal schooling from the British Group. The French (ETH1), the North Europeans (ETH3), the Native Indians (ETH8) and the other and unknown (ETH9) groups all have a significantly lower rate-of-return per year than the British. The returns per year of school for the Slavic (ETH4), the Italian (ETH5), the Jewish (ETH6) and the Oriental (ETH7) groups do not appear to be significantly different from the British Isles group.

Considering the experience-ethnicity interaction effects, it is apparent that the fffect of experience on earnings for three of the groups is significantly different than for the British Isles group. As indicated above, considering only the coefficients of EX and EE (hence the British Group) implies a peak in earnings at 32 years of experience. In comparison the earnings-experience profile for the French Group is significantly flatter and earnings peak at 33 years of experience. On the other hand, the earnings-experience profiles for the Slavic and Jewish groups appear to be significantly steeper than for the British group. In the case of the Slavic group, the coefficients imply that earnings peak at 27 years of experience. For the Jewish group, the profile is even steeper - the peak occurs at 33.4 years of experience, which is later than for either the French group of the British Isles group.

E.2.2. Mean Occupational Earnings

This is a continuous variable measuring the mean male earnings in each individual's occupation and province of residence (WAGE). It can be viewed as a proxy characterizing the regional levels of supply and demand for labor in an occupation, which make the occupation in which an individual is working either a high or low paying occupation in his or her province. As expected, the coefficient of this variable is positive and highly significant. In fact, if labor markets were perfect and all labor within an occupation was homogeneous, then all laborers should receive the occupational average wage rate. Thus, the WAGE variable and weeks worked variable would be the only significant variables in the regression. It is only because labor markets are not perfect and labor is heterogeneous that the other variables are significant.

E.2.3. Occupation

As a consequence of the heterogeneity of labor, workers within occupations earn different incomes based, in part, on the amount of human capital they possess. In some occupations, human capital is more important in determining incomes than in others (e.g., managerial-administrative versus clerical occupations). In those occupations where human capital is more important, it is reasonable to assume that workers will on the average, possess higher levels of human capital. As a consequence, having controlled for schooling and experiences, the inclusion of occupational dummies in the presence of the WAGE variable will result in negative coefficients for the dummy variables for all occupations in which human capital is more important in determining income than in clerical occupations which is the reference category.(4)

The occupational coefficients which were found to be significantly different from zero support this argument. Using a two-tailed test, occupations having negative coefficients significant at the 10% level were the following: managerial-administrative (OC1), natural science-engineering-mathematics (OC2), social science (OC3), religious (OC4), medical-health (OC6), and processing (OC13). These are the occupations in which either schooling and/or experience can be expected to have a strong positive effect on earnings (probably experience alone for processing). Occupations where the coefficients were negative but less significant - teaching (OC5), sales (OC9), and farming (OC11) - are also occupations in which experience can be expected to have a sizeable return, reinforced by a sizeable return to schooling for teaching. Only the "other occupations" category had a positive coefficient which was significant.

E.2.4. Industry

We had no a priori expectations about the partial effect on earnings of being in one of the ten industrial categories related to manufacturing (the reference category), after controlling for the effects of such variables as occupation, region, education, experience, hours and weeks worked, and the capital-to-labor ratio. Forestry (IN2) and construction (IN5) produce significantly higher earnings than manufacturing at the 10% significance level, and agriculture (IN1),

See footnote(s) on page 181.

fishing and trapping (IN3), trade (IN7), and community, business and personal services (IN9) seem to produce significantly lower earnings than manufacturing. The remaining industry coefficients are not significantly different from zero.

The positive differentials associated with the forestry and construction industries may be the result of union-maintained premiums to compensate for the highly seasonal nature of work in these industries. Employment in the agriculture and fishing and trapping industries is also highly seasonal, but in contrast, the forestry and construction industries are characterized by labor organizations which exercise a great deal of market power. (Note that variations in weeks of work are controlled for separately.)

E.2.5. Capital-to-labor Ratio

The industrial capital-to-labor ratio prevailing in a province was introduced to capture aspects of labor demand that depend on capital as a complementary factor. The availability of capital varies across industries and provinces. It will generally be the case that the higher the capital-to-labor ratio in an individual's employment setting, the higher will be his or her marginal reproductivity, and hence, the higher the demand for his or her services. Ceteris paribus, this should imply higher earnings. Therefore, the coefficient for KLR is expected to be positive. It turns out to be positive and significant at the 1% level with a one dollar increase in the capital per employee implying a 0.42% increase in annual earnings.

Since the capital intensiveness of a job may say something about its skill-learning content, there might be a significant interaction between the capital-to-labor ratio and the experience variables. It seems reasonable that highly-automated jobs, which would be associated with high KLRs would impart little in the way of skills, and hence, tend to have flatter earnings-experience profiles than less automated jobs. This in turn would imply a negative coefficient for the interaction for KLR and experience (EX·KLR) and a positive coefficient for the interaction of KLR and experience squared (EE·KLR). Both interaction coefficients have the expected signs and are significant at the 5% level, producing flatter earnings-experience profiles which peak later for all ethnic groups except Orientals and Native Indians.

E.2.6. Region

The regional coefficients for the Maritime Provinces (REG1), Quebec (REG2), Ontario (REG3) and the Prairie Provinces (REG4) are all negative and significant at the 1% level. Their magnitudes indicate that, *ceteris paribus*, earnings are highest in British Columbia, followed by Ontario, the Prairie Provinces, Quebec, and lastly, by the Maritime Provinces. This finding, which was expected, is noteworthy since it indicates that even after controlling for human capital factors, occupation, industry, size of community, regional differences in mean-occupational pay, regional differences in the levels of capital investment (measured by capital-to-labor ratios) and personal characteristics in general, the region of residence still has a significant effect on earnings. This implies that after one has controlled for differences in labor productivity, and hence, inter-personal differences in demand, as well as general differences in industrial and occupation demand, labor supply does not adjust to equalize nominal earnings.

E.2.7. Type of Community

A priori, we expect that the type of the community in which an individual resides will affect his or her earnings because of intra-regional differences in the specific job content of occupations and the nature of industrial activity. Note that at the intra-provincial level there is no control for variations in occupational mean wages or industrial capital-to-labor ratios. The estimated coefficients for place of residence in 1970 indicate that relative to urban areas of less than 30,000 (the control category), residence in urban areas of 30,000 or more (RES1) has a partial effect of producing significantly higher earnings, whereas residence in a rural non-farm community (RES2) produces significantly lower earnings. The negative but not significant coefficient for rural farm residence (RES3) indicates that farm earnings are not significantly lower than those of residents in urban areas of less than 30,000.

These findings are consistent with the results of many studies that have found a gross positive relationship between earnings and community size, and the prevalence of poverty in rural non-farm communities. These latter communities have been largely supplanted in their economic role of servicing farming by the urban areas of less than 30,000, and have become reservoirs of individuals who are either unable or unwilling to make the transition from rural to urban life.

E.2.8. Immigration and Migration

Migration, whether from foreign countries or within Canada, is likely to be selective with respect to age, education, and motivation, with the result that immigrants or migrants would tend to be younger, more educated, and more motivated than the average member of the population in the location of origin or destination. It is also likely to be the case that adaptation of the new environment takes time, particularly for the immigrant from a foreign culture. Hence, depending on the cultural gap involved, recent immigrants are likely to earn less than non-migrants or immigrants of older vintage. The fact that immigration or migration can be viewed as a human capital investment that will increase lifetime earnings relates to the individual's pre- and post-migration earnings, but has no clear implication for the earnings of migrants versus non-migrants, per se.

In terms of the variables used in this study, the net effect of these phenomena should be to make the coefficient for the variable representing post-1966 immigration from a foreign country (MIG) negative; for the variable representing foreign immigration between 1961-70 (PIM3), also negative; and the coefficients for earlier foreign immigration in the periods 1946-60 (PIM2), and before 1946 (PIM1) increasingly less negative, or even positive. Following this reasoning, the coefficient representing internal immigration since 1966 (MIG) should also be negative. None of these variables turned out to be significant at the 10% level. In general, this lack of significance indicates that after one has controlled for such things as ethnicity, occupation, industry, schooling, experience, and hours and weeks worked, recent internal migration and the period of immigration into Canada have no significant residual effects on earnings.

E.2.9. Marital Status

Being married as opposed to not being married is likely to have associated with it a variety of motivational factors related to the well-being of one's family. Thus, married males are expected to earn higher incomes than unmarried males and the coefficient for MS1 should be positive. Consistent with expectation, it is positive and significant at the 10% level.

E.2.10. Official Languages

Government policies, primarily at the federal level, to promote bilingualism may be expected to produce job opportunities favoring individuals who have a working knowledge of both English and French, which ceteris paribus, results in higher earnings. Insofar as individuals are aware of these opportunities and strive to master "the other official language", the knowledge of both official languages represents a dimension of motivation. This effect would be observed as a positive coefficient for LN02. This coefficient is positive, significant at the 5% level, and it implies an earnings premium of 3.2%.

E.2.11. Self-employment

Whether one is self-employed or not is likely to depend on another set of personal characteristics. The self-employed individual could be conceived of as being more adventurous, innovative, or prone to bearing risk; or simply more entrepreneurially inclined than his or her counterpart who works for a wage or salary. Furthermore, self-selection is likely to operate in such a way that those who fail on their own become employees. These factors suggest on balance that self-employment (SE), should have the partial effect of producing higher earnings. This is born out by a positive coefficient which is significant at the 1% level. The coefficient implies that, holding everything else constant, the earnings of self-employed individuals are 21% higher than their non-self-employed counterparts.

It also seems reasonable that the work experience of the self-employed is very different from experience of an employee; but, a priori, we have no particular expectation as to how the interactions of self-employment and experience should affect earnings. It turns out that both EX·SE and EE·SE are negative and significant at the 5% level. This implies, other things held constant, that self-employed individuals tend to have flatter experience-earnings profiles than people who are employees. Insofar as those who are now self-employed have been self-employed throughout most of their labor market involvement, the flatter profile may indicate that the self-employed learn by trial and error and have a fairly erratic earnings profile through time, as opposed to the employed individual who learns on-the-job from others and tends to steadily move up a job and salary hierarchy. Earnings peak substantially earlier for self-employed individuals. For the British, this peak occurs at 21 years of experience for the self-employed, as opposed to at 32 years for an employee working for a wage or salary.

E.2.12. Hours Usually Worked

Since annual earnings can be viewed as the product of an hourly wage rate, the number of hours usually worked per week, and the number of weeks worked per year, the effect of usually working less than 35 hours per week should be, ceteris paribus, to reduce annual earnings. Conversely, the effect of usually working 45 or more hours per week should be to increase annual earnings. The coefficient for the part-time hours variable (PTHRS) has the expected negative sign and is significant at the 1% level. However, even after controlling for the interaction of overtime work and self-employment (OTSE) which should capture the effect of the lower earnings received by the proprieters of small family enterprises, the coefficient for overtime work (OTHRS) is negative and insignificant. This indicates that beyond 44 hours per week, the additional hours of work do not increase earned income, possibly even decrease it. This suggests that overtime hours may be the result of an attempt to compensate, in terms of a target income level, for low hourly wage rates.

E.2.13. Weeks Worked

Variation in the number of weeks worked per year should have a positive association with annual earnings. Variation in weeks worked may be voluntary (a supply phenomenon) or involuntary (a demand phenomenon). These variations in the weeks of work are probably caused by many of the same factors which cause variations in the wage rate, as well as partly being a supply response to expected wage rates. This supply response effect poses a simultaneity problem which we have been unable to solve in a conceptually satisfactory way given the large number of dummy variables in the model. Nonetheless, we expect to find a strong positive partial relationship between the number of weeks worked and annual earnings even though the estimates of the coefficient of this variable may be inconsistent.

In order to determine if the implicit weekly wage rate is constant, increasing, or decreasing with respect to the number of weeks worked, we employed the logarithm of weeks worked (LWW) as an explanatory variable. Its coefficient can be interpreted as the elasticity of annual earnings with respect to a change in weeks worked. An elasticity (coefficient) equal to unity implies a constant implicit weekly wage rate, and an elasticity greater than unity indicates that the implicit weekly wage rate is an increasing function of weeks worked. The coefficient for LWW turned out to be 0.827 and significantly less than unity.[(5)] This indicates that earnings increase much less than proportionately with weeks worked - a 10% increase in the number of weeks worked only increases earnings by 8.3%. This may be due to the fact that the bulk of Canadian males who work less than a full year do so because of the seasonal character of their work and that, in order to attract labor to seasonal jobs, employers are forced to add a premium to the weekly wage for the weeks actually worked. Jointly, these would cause annual earnings to decrease (increase) less than proportionately with the decrease (increase) in weeks worked.

See footnote(s) on page 181.

FOOTNOTES

(1) For convenience, these latter two interactions were omitted from the computation of the rates-of-return to schooling discussed in Chapter 5.

(2) The comparable figures for the French group (ETH1) are $b_1 = 0.0270 - 0.0111 = 0.0159$, and $c_1 = -0.000425 + 0.000184 = -0.000241$. Thus, the French earnings profile will tend to rise more slowly and peak slightly later at 33 years of experience.

(3) This extremely large coefficient results from the very low returns to education and experience to the Native Indian group. In essence, an uneducated individual with no labor market experience will earn nearly as much as an educated, experienced worker in this group.

(4) The total effect of being in the ith occupation relative to clerical occupations is $a_i OC_i + b\,(WAGE_i)$, where a_i is the coefficient for the ith occupation's dummy, b is the coefficient for WAGE, and $WAGE_i$ is the across-province average of mean earnings in the ith occupation.

(5) Mincer (1974), using comparable census data for the United States, found the coefficient of the logarithm of weeks worked to be significantly greater than unity in a much simpler earnings function.

APPENDIX F

VARIANCE PARTITIONING PROCEDURE

This Appendix outlines the procedure used to partition the variance into its various components. The ith observation on Equation 3.6 can be written as

$$E_i = (Z_i\alpha + e_i) + (R_i\rho + H_iu_i) + (W_i\beta), \tag{F.1}$$

where Z_i is a 1xJ row vector of observations on the Z_{ij}'s, α is a Jx1 column vector of α_j's, R_i is a 1xK vector of observations on data (with typical elements S_i, S_iX_{ik}, $f(EX_i)\ X_{ik}$, etc.), ρ is a Kx1 column vector of parameters (with typical elements λ_k and γ_k), H_i is the human capital proxy defined in Equation 3.3, e_i and u_i are the error terms defined in Equations 3.4 and 3.5, W_i is a scalar observation on the logarithm of weeks worked, and β is a scalar coefficient. The three terms enclosed in parenthesis correspond to PME, HCE, and LSE, respectively.

Now suppose one individual is chosen at random. Thus, all the independent variables are random variables. The unconditional variance of E can be written as

$$V(E) = a'\Sigma_{ZZ}\,\alpha + \sigma_e^2 + \rho'\Sigma_{RR}\rho + (\sigma_H^2 + \mu_H^2)\sigma_u^2 + \beta'\Sigma_{WW}\,\beta$$

$$+ 2\alpha'\Sigma_{ZR}\,\rho + 2\,\alpha'\Sigma_{ZW}\,\beta + 2\beta'\Sigma_{WR}\,\lambda\,, \tag{F.2}$$

where Σ_{ZZ} and Σ_{RR} are the variance-covariance matrices of Z and R, respectively; Σ_{ZR} is the matrix of covariances between these vectors; Σ_{WW} denotes the scalar variance of W, and Σ_{ZW} and Σ_{ZR} denote the covariance matrices between W and Z and W and R; respectively; μ_H and σ_H^2 are the mean and variance of the level of human capital; all other symbols have been previously defined.

Starting with corresponding logarithm of the weekly wage rate equation,

$$LWR_i = (Z_i\alpha + e_i) + (R_i\rho + H_iu_i), \tag{F.3}$$

a parallel derivation produces the following expression for the unconditional variance of LWR

$$V(LWR) = \alpha'\Sigma_{ZZ}\alpha + \sigma_e^2 + \rho'\Sigma_{RR}\rho + (\sigma_H^2 + \mu_H^2)\,\sigma_u^2 + 2\alpha'\Sigma_{ZR}\rho. \tag{F.4}$$

The terms in F.2 and F.4 can be collected as follows:

$$V(PME) = \alpha' \Sigma_{ZZ} \alpha + \sigma_e^2;$$

$$V(HCE) = \rho' \Sigma_{RR} \rho + (\sigma_H^2 + \mu_H^2)\sigma_u^2;$$

$$V(LSE) = \beta' \Sigma_{WW} \beta;$$

$$C(PME, HCE) = \alpha' \Sigma_{ZR} \rho;$$

$$C(HCE, LSE) = \beta' \Sigma_{WR} \rho; \text{ and}$$

$$C(PME, LSE) = \alpha' \Sigma_{ZW} \beta.$$

PME denotes personal attributes and market effects; HCE denotes human capital effects; and LSE denotes labor supply effects. The terms in Equations F.2 and F.4 have been evaluated using maximum likelihood techniques and the resulting values are summarized in Tables 5.3 and 5.4.

APPENDIX G

COEFFICIENT ESTIMATES FOR THE FEMALES EARNINGS EQUATIONS

The first-and second-stage estimates of the full annual earnings model for females appear in Tables G.1 and G.2, respectively. Second-stage estimates for the corresponding female full wage rate model appear in Table G.3, and Table G.4 contains the second-stage estimates for the reduced annual earnings model. Each model was estimated separately for each of the five female ethnic subgroups. Since each model tells very much the same story, this appendix will concentrate exclusively on an interpretive discussion of the second-stage coefficient estimates obtained for the full annual earnings model.

G.1. Human Capital Effects

G.1.1. Formal Schooling

Each additional year of schooling is expected to increase annual earnings. As anticipated, the YSCH coefficient is positive for each ethnic grouping, and each is significant at the 1% level. The coefficient for the French females was highest, indicating a 4.0% increase in earnings per year of schooling. The East and South European grouping had the lowest coefficient; it indicated a 2.1% rate-of-return per year of schooling.

Having earned a university degree, as opposed to having attended university but not having earned a degree (DEG=1) is expected, *ceteris paribus*, to yield higher earnings. This variable turned out to have a positive effect for each ethnic grouping of females, and is significant at the 1% level for all except the Other female grouping (which was significant at the 10% level).

G.1.2. Experience

Earnings were expected to increase at a decreasing rate for each year of continuous work experience. The coefficients obtained for the potential experience variables adjusted for discontinuities in work experience (EXS=EX-HAT and EES $= EXS^2$) are consistent with this expectation. For each

TABLE G.1. First Stage Regression Results for the Full Annual Earnings Model for Females(1)

	Ethnic groups				
Variable	British	French	North European	East and South European	Other
Intercept	3.373 (0.112)	3.319 (.193)	3.149 (.207)	3.422 (.222)	3.683 (.267)
HH	.0912 (.0273)	.0831 (.0390)	.0533 (.0656)	.0850 (.0655)	.0360 (.0740)
MS1	.101 (.0171)	.108 (.0219)	.0910 (.0377)	.200 (.0363)	.0741 (.0460)
MS2	.0655 (.0249)	.0926 (.0374)	.135 (.0599)	.141 (.0594)	.0835 (.0742)
LNO2	-.0210 (.0236)	.0769 (.0174)	-.0318 (.0579)	-.0105 (.0394)	.00158 (.0553)
NCB	-.0573 (.00839)	-.0648 (.0111)	-.0461 (.0187)	-.0645 (.0204)	-.0591 (.0242)
NC2	.00391 (.00108)	.00522 (.00118)	-.00141 (.00247)	.00498 (.00302)	.00311 (.00311)
TM1	-.754 (.0139)	-.571 (.0209)	-.702 (.0303)	-.644 (.0290)	0.582 (.0415)
OTHRS	.0477 (.0236)	-.0723 (.0287)	0.114 (.0463)	0.0440 (.0454)	.0559 (.0556)
SE	-.222 (.0356)	-.143 (.0463)	-.311 (.0654)	-.0831 (.0648)	-.0684 (.0826)
DEG	.223 (.0308)	.184 (.0486)	.291 (.0716)	.225 (.0607)	.0981 (.0678)
OC1	-.202 (.0876)	-.326 (.170)	-.423 (.170)	-.380 (.189)	-.182 (.248)
OC2	-.119 (.0835)	.00133 (.147)	-.174 (.174)	.0823 (.147)	-.0839 (.182)
OC3	-.201 (.0604)	-.0490 (.0885)	-.163 (.145)	-.366 (.136)	.0840 (.166)
OC4	-.869 (.249)	-.381 (.315)	-.103 (.721)	.0457 (.494)	——(2)
OC5	-.00166 (.0643)	.111 (.111)	-.245 (.125)	-.145 (.134)	.113 (.162)

See footnote(s) at end of table.

TABLE G.1. First Stage Regression Results for the Full Annual Earnings Model for Females(1) (Continued)

	Ethnic groups				
Variable	British	French	North European	East and South European	Other
OC6	.0978 (.0296)	.0510 (.0631)	.0536 (.0601)	-.00905 (.0668)	.0430 (.0827)
OC7	.0681 (.0626)	.0072 (.0882)	-.180 (.142)	.0961 (.122)	.266 (.193)
OC8	-.0260 (.0283)	.0270 (.0439)	-.0751 (.0591)	-.0627 (.0561)	-.181 (.0703)
OC9	.0105 (.0386)	-.0907 (.0646)	.0835 (.0731)	-.0663 (.0775)	-.121 (.108)
OC10	-.102 (.0455)	-.0422 (.0774)	.0640 (.0829)	.00681 (.0900)	-.0687 (.111)
OC11	-.343 (.0594)	-.212 (.123)	-.131 (.0930)	-.218 (.105)	-.137 (.137)
OC12	.396 (.199)	.0301 (.228)	.283 (.519)	.317 (.702)	.774 (.366)
OC13	.107 (.0530)	.147 (.0621)	.299 (.0964)	.124 (.0791)	-.168 (.106)
OC14	.0314 (.0367)	.0658 (.0534)	.112 (.0695)	-.111 (.0555)	-.123 (.0706)
OC15	.0502 (.132)	-.0912 (.196)	.532 (.362)	.800 (.249)	-1.470 (.685)
OC16	-.103 (.0998)	.505 (.223)	.221 (.198)	.771 (.289)	.0719 (.350)
RES1	.0634 (.0154)	.0412 (.0208)	.0182 (.0335)	.0985 (.0402)	.0790 (.0558)
RES2	0.0668 (.0210)	-.0613 (.0298)	-.0330 (.0450)	-.0170 (.0660)	-.162 (.0774)
RES3	-.0747 (.0358)	-.219 (.0542)	-.0930 (.0610)	-.0652 (.0921)	-.291 (.139)
REG1	-.0873 (.0301)	.00414 (.0793)	-.0654 (.0794)	.222 (.144)	.0531 (.121)
REG2	-.0173 (.0299)	.0173 (.0629)	.105 (.0760)	.0193 (.0538)	-.0685 (.0634)

See footnote(s) at end of table.

TABLE G.1. First Stage Regression Results for the Full Annual Earnings Model for Females(1) (Concluded)

	Ethnic groups				
Variable	British	French	North European	East and South European	Other
REG3	-.0187 (.0184)	-.0565 (.0649)	-.0634 (.0388)	.00314 (.0448)	-.0849 (.0511)
REG4	-.0260 (.0229)	.0121 (.0738)	-.0575 (.0398)	-.00437 (.0480)	-.122 (.0606)
YSCH	-.0373 (.00290)	.0391 (.00382)	.0361 (.00617)	.0201 (.00487)	.0231 (.00584)
EX	.0254 (.00142)	.0259 (.00211)	.0290 (.00308)	.0218 (.00290)	.0249 (.00379)
EE	-.000351 (.0000270)	-.000379 (.0000400)	-.000402 (.0000574)	-.000260 (.0000530)	-.000308 (.0000684)
WAGE	.000162 (.0000296)	.000183 (.0000520)	.000241 (.0000538)	.000209 (.0000589)	.000171 (.0000705)
LWW	.943 (.00895)	.908 (.0129)	.951 (.0192)	.915 (.0188)	.917 (.0257)
R^2	.665	.608	.658	.615	.624
N	15,093	7,623	3,298	3,592	1,875

(1) Numbers in parentheses are estimated standard errors.

(2) There were no observations in this cell.

Source: Public Use Sample Tape.

TABLE G.2. Second Stage Regression Results for the Full Annual Earnings Model for Females[1]

	Ethnic group				
Variable	British	French	North European	East and South European	Other
Deg	.2202 (.031)	.1769 (.049)	.2841 (.072)	.2163 (.061)	.09726 (.068)
OC 1	-.2001 (.088)	-.3308 (.17)	-.4241 (.17)	-.3741 (.19)	-.1620 (.25)
OC 2	-.1216 (.084)	-.006878 (.15)	-.1756 (.17)	.08119 (.15)	-.07651 (.18)
OC 3	-.2039 (.060)	-.05291 (.088)	-.1702 (.15)	-.3702 (.14)	.08194 (.17)
OC 4	-.8688 (.25)	-.3638 (.32)	-.1318 (.72)	.05237 (.49)	——[2]
OC 5	-.002521 (.064)	.1109 (.11)	-.2521 (.12)	-.1433 (.13)	.1152 (.16)
OC 6	.09774 (.030)	.05028 (.063)	.05387 (.060)	-.007759 (.067)	.04822 (.083)
OC 7	.06853 (.063)	.07877 (.088)	-.1837 (.14)	.1002 (.12)	.2645 (.19)
OC 8	-.02915 (.028)	.03038 (.044)	-.07923 (.059)	-.06287 (.056)	-.1825 (.070)
OC 9	.006719 (.039)	-.08765 (.065)	.08553 (.073)	-.06592 (.078)	-.1209 (.11)
OC 10	-.1059 (.045)	-.04008 (.077)	.06325 (.083)	.005183 (.090)	-.07214 (.11)
OC 11	-.3469 (.059)	-.2117 (.12)	-.1325 (.093)	-.2160 (.10)	-.1358 (.14)
OC 12	.3960 (.20)	.03215 (.23)	.2990 (.52)	.2966 (.70)	.7762 (.37)

See footnote(s) at end of table.

TABLE G.2. Second Stage Regression Results for the Full Annual Earnings Model for Females[1] (Continued)

	Ethnic group				
Variable	British	French	North European	East and South European	Other
OC 13	.1059 (.053)	.1531 (.062)	.3055 (.096)	.1277 (.079)	-.1719 (.11)
OC 14	.03217 (.037)	.06927 (.053)	.1139 (.070)	-.1058 (.056)	-.1241 (.071)
OC 15	.05150 (.13)	-.08994 (.20)	.5399 (.36)	.7952 (.25)	-1.422 (.68)
OC 16	-.1029 (.10)	.5031 (.22)	.2272 (.20)	.7840 (.29)	.09071 (.35)
MS 1	.1164 (.017)	.1215 (.022)	.1146 (.035)	.2172 (.035)	.09012 (.044)
MS 2	.07219 (.025)	.09545 (.037)	.1464 (.059)	.1504 (.059)	.09207 (.074)
LNO 2	-.02542 (.024)	.07742 (.017)	-.03656 (.058)	-.01311 (.039)	.004046 (.055)
SE	-.2189 (.036)	-.1414 (.046)	-.3049 (.065)	-.08283 (.065)	-.06949 (.082)
RES 1	.06266 (.015)	.04059 (.021)	.01894 (.034)	.1023 (.040)	.07608 (.056)
RES 2	-.06712 (.021)	-.06181 (.030)	-.03087 (.045)	-.01731 (.066)	-.1601 (.077)
RES 3	-.07880 (.036)	-.2195 (.054)	-.09295 (.061)	-.06811 (.092)	-.2950 (.14)
REG 1	-.08648 (.030)	.003504 (.079)	-.06386 (.079)	.2219 (.14)	-.05598 (.12)
REG 2	-.01508 (.030)	.01583 (.063)	-.1016 (.076)	.02110 (.054)	-.06635 (.063)
REG 3	-.01788 (.018)	.05750 (.065)	-.06212 (.039)	.003234 (.045)	-.08513 (.051)

See footnote(s) at end of table.

TABLE G.2. Second Stage Regression Results for the Full Annual Earnings Model for Females(1) (Concluded)

Variable	Ethnic group: British	French	North European	East and South European	Other
REG 4	-.02593 (.023)	.01362 (.074)	-.05676 (.040)	-.005796 (.048)	-.1218 (.061)
TM1	-.7523 (.014)	-.5692 (.021)	-.6968 (.030)	-.6432 (.029)	-.5807 (.041)
HH	.09989 (.027)	.08976 (.039)	.06547 (.065)	.09144 (.065)	.03723 (.074)
WAGE	.0001618 (.00003)	.0001855 (.00005)	.0002437 (.00005)	.0002081 (.00006)	.0001675 (.00007)
YSCH	.03775 (.0029)	.03994 (.0038)	.03684 (.0062)	.02097 (.0049)	.02359 (.0058)
LWW	.9424 (.0090)	.9056 (.013)	.9494 (.019)	.9156 (.019)	.9161 (.026)
OTHRS	.04727 (.024)	-.06907 (.029)	-.1142 (.046)	-.04595 (.045)	.05796 (.056)
EXS	.02471 (.0014)	.02717 (.0022)	.02920 (.0032)	.02083 (.0029)	.02580 (.0039)
EES	-.0003629 (.00003)	-.0004312 (.00004)	-.0004398 (.00006)	-.0002646 (.00006)	-.0003591 (.00008)
HAT	-.01977 (.0032)	-.02300 (.0045)	-.02583 (.0052)	-.01486 (.0057)	-.01799 (.0076)
INTERCEPT	3.3816	3.3102	3.1426	3.4394	3.6942
R^2	.66	.61	.66	.61	.62
N	15,093	7,623	3,298	3,592	1,875

(1) Numbers in parentheses are estimated standard errors.

(2) There were no observations in this cell.

Source: Public Use Sample Tape.

TABLE G.3. Second Stage Regression Results for the Full Wage Rate Model for Females(1)

Variable	Ethnic group				
	British	French	North European	East and South European	Other
DEG	.2217 (.031)	.1831 (.049)	.2812 (.072)	.2149 (.061)	.1023 (.068)
OC 1	-.2084 (.088)	-.3057 (.17)	-.4228 (.17)	-.3838 (.19)	-.1698 (.25)
OC 2	-.1243 (.084)	.009957 (.15)	-.1692 (.17)	.09214 (.15)	-.07757 (.18)
OC 3	-.1954 (.061)	-.01683 (.089)	-.1624 (.15)	-.3472 (.14)	.06859 (.17)
OC 4	-.8693 (.25)	-.3494 (.32)	-.1485 (.72)	.1025 (.50)	___(2)
OC 5	-.009059 (.064)	.1248 (.11)	-.2465 (.13)	-.1557 (.13)	.1188 (.16)
OC 6	.09691 (.030)	.06310 (.063)	.05822 (.060)	-.01363 (.067)	.04666 (.083)
OC 7	.07176 (.063)	.1046 (.088)	-.1801 (.14)	.1328 (.12)	.2737 (.19)
OC 8	-.02092 (.028)	.03977 (.044)	-.07354 (.059)	-.04561 (.056)	-.1823 (.070)
OC 9	.01372 (.039)	-.08210 (.065)	.08740 (.073)	-.05303 (.078	-.1240 (.11)
OC 10	-.09129 (.046)	-.03316 (.078)	.06807 (.083)	.02393 (.090)	-.07585 (.11)
OC 11	-.3278 (.059)	-.1776 (.12)	-.1188 (.093)	-.1821 (.10)	-.1077 (.14)
OC 12	.4440 (.20)	.09220 (.23)	.2985 (.52)	.3044 (.70)	.8183 (.37)
OC 13	.1272 (.053)	.1684 (.062)	.3181 (.096)	.1590 (.079)	-.1618 (.11)

See footnote(s) at end of table.

TABLE G.3. Second Stage Regression Results for the Full Wage Rate Model for Females[1] (Continued)

	Ethnic group				
Variable	British	French	North European	East and South European	Other
OC 14	.04284 (.037)	.07704 (.054)	.1203 (.070)	-.08694 (.056)	-.1226 (.071)
OC 15	.06224 (.13)	-.07692 (.20)	.5545 (.36)	.8366 (.25)	-1.3066 (.69)
OC 16	-.1005 (.10)	.5242 (.22)	.2344 (.20)	.7993 (.29)	.09977 (.35)
MS 1	.1117 (.017)	.1207 (.022)	.1104 (.035)	.2098 (.035)	.08926 (.044)
MS 2	.06616 (.025)	.09100 (.038)	.1360 (.060)	.1443 (.059)	.09566 (.074)
LNO 2	-.02316 (.024)	.07863 (.017)	-.02729 (.058)	-.005885 (.039)	.003185 (.055)
SE	-.2267 (.036)	-.1598 (.046)	-.3184 (.065)	-.09319 (.065)	-.07389 (.083)
RES 1	.06129 (.015)	.04017 (.021)	.01610 (.034)	.09835 (.040)	.06259 (.056)
RES 2	-.06583 (.021)	-.05575 (.030)	-.03019 (.045)	-.009557 (.066)	-.1563 (.077)
RES 3	-.07613 (.036)	-.2168 (.054)	-.09391 (.061)	-.07088 (.092)	-.2997 (.14)
REG 1	-.08421 (.030)	-.002848 (.080)	-.06737 (.079)	.2263 (.14)	-.05202 (.12)
REG 2	-.01540 (.030)	.008785 (.063)	-.1078 (.076)	.01330 (.054)	-.06033 (.063)
REG 3	-.01961 (.018)	.04847 (.065)	-.06347 (.039)	-.001491 (.045)	-.08626 (.051)
REG 4	-.02543 (.023)	-.0007252 (.074)	-.05909 (.040)	-.01109 (.048)	-.1267 (.061)
TM1	-.7259 (.013)	-.5229 (.020)	-.6729 (.029)	-.6058 (.028)	-.5449 (.040)

See footnote(s) at end of table.

TABLE G.3. Second Stage Regression Results for the Full Wage Rate Model for Females[1] (Concluded)

	Ethnic group				
Variable	British	French	North European	East and South European	Other
HH	.09849 (.027)	.08992 (.039)	.06460 (.066)	.08165 (.066)	.03118 (.074)
WAGE	.0001638 (.00003)	.0001769 (.00005)	.0002412 (.00005)	.0002120 (.00006)	.0001627 (.00007)
YSCH	.03746 (.0029)	.03906 (.0038)	.03676 (.0062)	.02111 (.0049)	.02329 (.0059)
OTHRS	.05048 (.024)	-.06760 (.029)	-.1121 (.047)	-.04922 (.046)	.05601 (.056)
EES	-.0003372 (.00003)	-.0003785 (.00004)	-.0004159 (.00006)	-.0002222 (.00006)	-.0003284 (.00008)
HAT	-.01846 (.0032)	-.02025 (.0046)	-.02474 (.0051)	-.01268 (.0056)	-.01646 (.0075)
EXS	.02292 (.0014)	.02368 (.0021)	.02767 (.0031)	.01802 (.0029)	.02348 (.0038)
INTERCEPT	3.1879	3.03587	2.9864	3.1505	3.4449
R^2	.33	.26	.32	.25	.28
N	15,093	7,623	3,298	3,592	1,875

(1) Numbers in parentheses are estimated standard errors.

(2) There were no observations in this cell.

Source: Public Use Sample Tape.

TABLE G.4. Second Stage Regression Results for the Reduced Annual Earnings Model for Females[1]

	Ethnic group				
Variable	British	French	North European	East and South European	Other
DEG	.2382 (.030)	.2324 (.049)	.3063 (.069)	.2699 (.060)	.2041 (.067)
RES 1	.06140 (.016)	.04961 (.021)	.02866 (.034)	.1047 (.041)	.08697 (.057)
RES 2	-.08718 (.022)	-.05708 (.031)	-.05256 (.046)	-.01557 (.068)	-.08568 (.078)
RES 3	-.1239 (.036)	-.2831 (.054)	-.1883 (.059)	-.1496 (.085)	-.3649 (.14)
REG 1	-.1444 (.024)	-.02472 (.074)	-.1713 (.076)	.1752 (.14)	-.1176 (.12)
REG 2	.07434 (.028)	.1368 (.064)	-.004566 (.072)	.07579 (.051)	.004667 (.059)
REG 3	.04328 (.018)	.1514 (.067)	.007524 (.038)	.05471 (.044)	-.01455 (.049)
REG 4	-.05410 (.022)	.01213 (.075)	-.1034 (.038)	-.04338 (.046)	-.1275 (.060)
TM1	-.7933 (.014)	-.5759 (.021)	-.7196 (.030)	-.6784 (.029)	-.5965 (.041)
EM	.1523 (.017)	.1473 (.022)	.1605 (.036)	.2576 (.036)	.1028 (.045)
YSCH	.06629 (.0027)	.07611 (.0035)	.06188 (.0058)	.04180 (.0043)	.05103 (.0052)
HAT	-.01795 (.0027)	-.02341 (.0046)	-.02545 (.0047)	-.01435 (.0052)	-.01865 (.0078)
EXS	.02714 (.0015)	.03196 (.0023)	.03258 (.0032)	.02083 (.0030)	.02847 (.0039)
EES	-.0004250 (.00003)	-.0005166 (.00005)	-.0005335 (.00007)	-.0003007 (.00006)	-.0004060 (.00008)

See footnote(s) at end of table.

TABLE G.4. Second Stage Regression Results for the Reduced Annual Earnings Model for Females[1] (Concluded)

	Ethnic group				
Variable	British	French	North European	East and South European	Other
LWW	.9645 (.0092)	.9319 (.013)	.9603 (.020)	.9317 (.019)	.9353 (.026)
INTERCEPT	3.446	3.343	3.533	3.750	3.727
R^2	.64	.57	.63	.59	.59
N	15,093	7,623	3,298	3,592	1,875

(1) Numbers in parentheses are estimated standard errors.

Source: Public Use Sample Tape.

ethnic grouping, the coefficient for EXS was positive and the coefficient for EES was negative and much smaller in absolute size. Each coefficient was highly significant for each ethnic grouping. The effect of initial years of experience are greatest for the North European females (2.9% per year) and least for the East and South European females (2.1% per year). The rate of decline in the return to experience depends on the level of experience, and this rate of decline is largest for the North European group (.088 EXS % per year), whereas for the East and South European females it is the smallest (.052 EXS % per year). Together, these two effects imply an earnings-experience profile for continuously working North European females which rises steeply at first but levels out relatively quickly. For the East and South European group, the profile rises more slowly initially but also levels off at a slower rate. For these two groupings of females, the coefficients imply maximum earnings being reached, respectively, at 33 and 39 years of continuous work experience. For British females maximum earnings would be reached at 34 years of continuous experience, for French females at 31 years, and for Other females at 36 years.

When an individual withdraws from labour market, for purposes other than to engage in formal schooling, human capital accumulation is presumed to cease and consequently earnings potential cannot increase. Work related skills can be expected to deteriorate from lack of use, resulting in a depreciation of the individual's already acquired human capital stock. This will result in the individual's earnings potential declining with the length of the hiatus from work activity. Thus we expect to find annual earnings negatively related to the length of hiatus from work experience computed using first-stage results (see Appendix C). The length of hiatus variable ($\hat{HAT}$) turns out to have been negative and highly significant coefficient for each of the ethnic groupings. In absolute size the coefficient is largest for North European females, each year of absence from work activity due to childbearing and rearing implying a 2.6% decline in annual earnings. East and South European females have the smallest coefficient, with each year of absence from work activity implying a 1.5% decline in earnings. The fact that the coefficient for North European females greatly exceeds the coefficient for the East and South European females cannot simply be interpreted as a more rapid depreciation of North European females' human capital, since the earnings increase that would have resulted from an additional year of experience would have been larger for the North European than East or South European female.

G.2. Personal Attributes and Market Effects

G.2.1. Marital Status

On the basis of the discussion in Chapter 4 we expected that being currently married (MS1=1), all other things constant, would tend to reduce earnings, whereas being widowed, divorced or separated (MS2=1) would tend to increase earnings. Our expectations were fulfilled for the latter variable: the coefficient for MS2 was positive for each female ethnic grouping and significant in each instance except for the Other females. Unexpectedly, the coefficients for MS1 turned out to be significantly positive for every ethnic grouping. The explanation for the surprising result may be self-selection bias. The women who work when they are married, and hence appear in our sample of married women, according to basic labour supply theory are only those women whose market wages exceed their reservation wage (determined by the value of their time spent in household activities). To the extent that the married women in our sample tend to be those who, given otherwise identical attributes, have higher earnings capacities and not those with lower household productivity, MS1 is serving as a proxy for otherwise unexplained sources of higher earnings and not the motivational effects of marriage on earnings. At any rate, the positive effect of marriage on female earnings, ceteris paribus, is consistent with the findings of Oaxaca (1973) with U.S. data and Gunderson (1976) with Canadian data.

G.2.2. Family Head

Being the head of a family unit is usually associated with a single female being financially independent, or else a female being the primary breadwinner of a multi-person family. Either situation would lead us to expect the coefficient for being a family head (HH=1) to be positive. All of these coefficients turned out to have a positive sign, but they were statistically significant only for the British and French groups.

G.2.3. Being Bilingual

Knowledge of both official languages was expected, ceteris paribus, to result in higher earnings. This only turned out to be the case for French females. The coefficient for LNO2 was positive and highly significant for this group, whereas it was not significantly different from zero for any of the other groups.

G.2.4. Self-employment

Using the same line of argument applied to the male model (see Appendix E), we expect self-employment to have a positive effect on female earnings. The coefficient for SE turned out to be significantly different from zero for only the British, the French and the North European groups. However, in each instance the coefficient was negative and large, implying that self-employed females earned from 14% to 30% less, ceteris paribus, than their counterparts working for wages and salaries. For these three groups, at least, self-employment may represent a desire for independence (and possibly flexible hours in the case of married women) which must be purchased at the price of much lower earnings.

G.2.5. Hours Usually Worked per Week

Usually working fewer than 35 hours per week, was expected to reduce annual earnings. This expectation was fulfilled as each ethnic grouping of females had a large negative coefficient for TM1 that was significant at the 1% level. On the other hand, usually working more than 44 hours per week was expected to increase annual earnings. The British group, however, was the only one for which the coefficient of OTHRS was positive and significant (at the 5% level). For the French and North European females, the OTHRS coefficient was negative and significant at the 1% level. The East and South European females and the Other females had coefficients which were not significantly different from zero. Thus, it appears that for all of the female groups, except the British, women work overtime hours where they are attempting to achieve some target level of income in a job with a low wage rate.

G.2.6. Weeks Worked

Annual earnings are expected to increase with the number of weeks worked. Thus we expect the logarithm of weeks worked (LWW) to have a positive coefficient. Each ethnic grouping of females turned out to have a highly significant positive coefficient for this variable. These coefficients range in size from .95 for the North European females to .91 for the French females, indicating that earnings increase proportionately less than weeks of work.

G.2.7. Mean Occupational Earnings

WAGE is the mean-female earnings in the province of residence for the occupation in which a female works. It is used to control for interregional variations in the labor supply-demand situation within occupations. If a woman works in what is generally a high paying occupation within her province of residence, it is anticipated that, ceteris paribus, her annual earnings will be higher than if she did not. Thus, the WAGE coefficient is expected to be positive. This turned out to be the case for the annual earnings equations estimated for each grouping of females; each had a positive WAGE coefficient that was highly significant.

G.2.8. Occupation

Most of the occupational coefficients for each grouping of females, except British females, were not significantly different from zero. This indicates that after one has controlled for acquired human capital, personal attributes, mean provincial-occupational earnings and location, relative to the clerical occupations, being in most other occupations does not produce significantly different earnings for females. Whereas we had no particular ex ante expectation about the signs of individual occupational coefficients, the reasoning used in the discussion of the occupational coefficients for males (see Appendix E) suggests that occupations whose coefficients turn out to be significantly negative are those in which human capital is relatively more important in determining earnings than in the clerical occupations, with the converse being true for those occupations whose coefficients turn out to be significantly positive. For the British females, this would imply that managerial-administrative (OC1), social science (OC3), religion related (OC4), service (OC10) and agricultural (OC11) occupations are the ones for that ethnic group in which human capital is a more important determinant of earnings than it is in the clerical occupations, whereas in the medical-health (OC6), other primary (OC12), and processing (OC13) occupations, human capital is a less important determinant of earnings than it is in clerical occupations. In this context the highly significant negative coefficient for agricultural occupations and the highly significant positive coefficient for medical-health occupations appear to be anomalies. It may, however, be the case that within these two broad occupational categories women are concentrated in jobs in which human capital is correspondingly more and less important than it is in the clerical occupations.

The signs of significant coefficients for the same occupational category are generally consistent across ethnic groupings of females. The only sign reversals occur for the machining-fabrication-assembling-repairing occupations (OC14) and the construction trades occupations (OC15). These are particularly broad occupational categories in which ethnic groups may concentrate in distinct types of jobs offering very different rewards for the accumulation of human capital.

G.2.9. Type of Community

Across ethnic groups, the magnitudes of coefficients for the place of residence variables indicate that earnings tend to be higher, ceteris paribus, the larger the size of the community in which a female resides. The reference category is urban areas with a population of less than 30,000. The coefficients for urban areas greater than 30,000 are positive, while those for rural nonfarm and farm residences are increasingly negative.

G.2.10. Region

The coefficients for the regional control variables will be statistically significant only if after controlling for all other factors such as level of education, weeks of work, provincial occupational wage, etc., there was still an unexplained residual associated with earnings levels in a particular region relative to British Columbia (the excluded category). It turned out that except for a negative Maritime Provinces (REG1) coefficient for the British females and negative Ontario (REG3) and Prairie Provinces (REG4) coefficients for Other females, none of the regional coefficients were significantly different from zero. Therefore, one may presume that, in general, geographic factors do not produce residual interregional earnings differentials (relative to British Columbia) for females. That is, after one has controlled for schooling, experience, labor supplied, occupation, size of community, etc., the part of Canada a woman lives in has very little effect on her earnings.

APPENDIX H

PARTITIONING THE EARNINGS DIFFERENTIAL BETWEEN TWO GROUPS

This appendix outlines the procedure used to partition the earnings differential between two groups of individuals into that portion which can be explained by differences in attribute levels and that portion which results from differences in rates-of-return on attributes (which has sometimes been attributed to discrimination).

Assume that the earnings functions have been correctly specified and estimated without bias for two separate groups using ordinary least squares. Then, for each ethnic group,

$$\overline{E}^s = \sum_{i=1}^{J} b_i^s \overline{X}_i^s, \qquad s = 1, 2, \qquad (H.1)$$

where $\overline{E}^s$ and $\overline{X}_i^s$ denote, respectively, the sample means of the logarithm of annual earnings and the ith attribute for the respective subgroup, and the b_i^s denote estimated parameters. The logarithmic mean earnings differentials can be expressed either as

$$\overline{E}^1 - \overline{E}^2 = \sum_i b_i^1(\overline{X}_i^1 - \overline{X}_i^2) + \sum_i \overline{X}_i^1(b_i^1 - b_i^2) - \sum_i (\overline{X}_i^1 - \overline{X}_i^2)(b_i^1 - b_i^2) \qquad (H.2)$$

or as

$$\overline{E}^1 - \overline{E}^2 = \sum_i b_i^2(\overline{X}_i^1 - \overline{X}_i^2) + \sum_i \overline{X}_i^2(b_i^1 - b_i^2) + \sum_i (\overline{X}_i^1 - \overline{X}_i^2)(b_i^1 - b_i^2). \qquad (H.3)$$

Equation H.2 and H.3 each partition the mean logarithmic earnings differences between the two groups into three components. The first terms of the right hand sides of these equations, $\sum_i b_i^s(\overline{X}_i^1 - \overline{X}_i^2)$, s = 1,2, are alternative measures of the difference in earnings that is accounted for by differences in the mean levels of attributes using the regression weights of Groups 1 and 2, respectively.

Masters (1974) considers the middle terms on the RHS, $\sum_i \overline{X}_i^s (b_i^1 - b_i^2)$, s= 1,2, to be alternative measures of discrimination using attribute levels of Groups 1 and 2, respectively, as weights. Oaxaca (1973), on the other hand, defines discrimination as the gross income differential less the amount "explained" by mean attribute differences. Thus, based on Equation H.2, Oaxaca's definition of discrimination is

$$\bar{E}^1 - \bar{E}^2 - \sum_i b_i^1(\bar{X}_i^1 - \bar{X}_i^2) = \sum_i \bar{X}_i^1(b_i^1 - b_i^2) - \sum_i(\bar{X}_i^1 - \bar{X}_i^2)(b_i^1 - b_i^2). \quad \text{(H.4)}$$

Expanding the last term on the right-hand side of the Equation H.4 and collecting terms yields

$$\sum_i \bar{X}_i^1(b_i^1 - b_i^2) - \sum_i(\bar{X}_i^1 - \bar{X}_i^2)(b_i^1 - b_i^2) = \sum_i \bar{X}_i^2(b_i^1 - b_i^2), \quad \text{(H.5)}$$

which is precisely equal to Masters' definition based on the regression coefficients from Group 2. Similarly it can be shown that Oaxaca's measure based on Equation H.3 is equivalent to Masters' definition using Group 1 regression coefficients.

Oaxaca's and Masters' procedures each yield two measures because of the inherent index-number problem involved in choosing the appropriate weights, and the two measures differ by the magnitude of the interaction term, $\sum_i (\bar{X}_i^1 - \bar{X}_i^2)(b_i^1 - b_i^2)$. Sufficient but not necessary conditions for this interaction term to be zero are that either $b_i^1 = b_i^2$ for all i (i.e., the rates-of-return are the same between the two groups and all earnings differences are explained by differences in attributes levels) or $\bar{X}_i^1 = \bar{X}_i^2$ for all (i.e., attribute levels are the same for the two groups). In the latter case, there is no problem of choosing the appropriate weights. Thus, a nonzero interaction term results from an index number problem of differing attribute levels and from an index number problem of differing attribute levels and differing rates-of-return between the two groups. Consequently, it seems reasonable that only part, but not all, of this interaction term should be included in the measure of the rates-of-return or "discrimination" effects. Hence, Oaxaca's and Masters' two measures should be considered upper and lower bounds on the extent of "discrimination". Furthermore, if b_i^1 and b_i^2 differ for reasons other than discrimination (such as motivation or self-selection), the sum of the attribute-weighted differences in the rates-of-return will not be pure measures of discrimination, but will be confounded by these other factors. For this reason, we prefer to use the term "residual" or "rates-of-return" effects rather than discrimination. This is particularly true in view of some of the findings reported in Chapter 7.

APPENDIX I

SPECIFICATION OF THE EARNINGS FUNCTIONS USED IN CHAPTER 7

Observations were confined to individuals age 30 and over who reported positive wage or salary earnings and had no property or non-wage income in 1970. The resulting sample sizes are indicated in Table I.1.

For each ethnic - sex - marital status group except ever-married females, a basic model specification was estimated which had the logarithm of wages and salaries (LWS) as the dependent variable. The independent variables in this basic or reduced model included the following:(1) YSCH, EX, EE, LWW, SQ, SQD, COL, TM1, DEG, IMIG(2), MIG, LNO2, ASCH, and an intercept. In addition, groups of dummy variables were included to control for region, community type, and period of immigration.(3) An expanded version of this specification, the full model, was also estimated which included controls for industry and occupation of employment. Since the sample sizes in some cases tended to be quite small (see Table I.1), it was necessary to aggregate some of the industrial and occupational variables as follows: OC2 + OC3; OC4 + OC5; OC11 + OC12; OC13 + OC14 + OC15 + OC16; and IN1 + IN2 + IN3 + IN4. This resulted in 10 dummy variables for occupation and eight dummy variables for industry.

Since no reasonable proxy for labor force experience was available for ever-married women, the model specification for ever-married women had to be altered slightly. For the ever-married women groups, AGE and AGE squared were used instead of EX and EE. In addition, the number of children ever-born (NCB) and the number of children ever-born squared (NC2) were also included. This specification is called the "female specification". In order to make the comparisons between ever-married females and other groups as discussed in Chapter 7 and Appendix H, it was necessary to also estimate the female specification for the comparison groups (never-married females and all males).

See footnote(s) on page 207.

TABLE I.1. Sample Sizes

Ethnic group[1]	Never-married		Ever-married	
	Male	Female	Male	Female
British	1,015	928	12,379	7,249
French	723	723	7,139	2,664
North European	237	165	3,037	1,571
East and South European	272	162	3,499	1,780
Other	192	139	1,746	796

(1) These groupings correspond to those discussed in Chapter 6.

Source: Public Use Sample Tape.

FOOTNOTES

(1) The variables and groups of variables are all defined in Appendix D.

(2) IMIG was disaggregated to separately account for interprovincial and intercounty moves.

(3) This was aggregated into two categories, PIM3 and other.

BIBLIOGRAPHY

Aitchison, J., and Brown, J. A. C. 1969. The Lognormal Distribution. Cambridge: Cambridge University Press.

Atkinson, A. B. 1970. On the Measurement of Inequality. Journal of Economic Theory 2: pp. 244-63.

Becker, G. S. 1964. Human Capital. New York: Columbia University Press.

Becker, G. S. 1967. Human Capital and the Personal Distribution of Income. Ann Arbor: Woytinsky Lecture, University of Michigan.

Becker, G. S. 1975. Human Capital. 2nd ed. New York: Columbia University Press.

Becker, G. S. 1976. The Economic Approach to Human Behavior. Chicago: The University of Chicago Press.

Ben-Porath, Y. 1967. The Production of Human Capital and the Life Cycle of Earnings. Journal of Political Economy 75: pp. 352-65.

Blinder, A. S. 1973. Wage Discrimination: Reduced Form and Structural Estimates. Journal of Human Resources 8: pp. 436-55.

Blinder, A. S. 1974. Toward an Economic Theory of Income Distribution. Cambridge: MIT Press.

Blinder, A. S., and Weiss, Y. 1974. Human Capital and Labor Supply: a Synthesis. Princeton University, Industrial Relations Section Working Paper No. 55.

Boulet, J.-A., and Raynauld, A. L'analyse des disparités de revenus suivant l'origine ethnique et la langue sur le marché montréalais en 1961. Economic Council of Canada. Discussion Paper No. 83.

Brown, G. D. 1978. Discrimination and Pay Disparities Between White Men and Women. Monthly Labor Review 101 (No. 3): pp. 17-22.

Canada. Statistics Canada. 1971. 1971 Census of Canada, Occupation of Classification Manual. Vol. 1, Catalogue No. CS 12-536, Ottawa: Information Canada.

Canada. Dominion Bureau of Statistics. 1972. Econometric Study of Incomes of Canadian Families, 1967. Catalogue No. 13-537 occasional. Ottawa: Information Canada.

Canada. Statistics Canada. 1975. 1971 Census of Canada, Incomes of Canadians, Employment Income by Sex, Occupation and Schooling for Provinces. Catalogue No. 94-769, Vol. 3 - Part 6 (Bulletin 3. 6-11). Ottawa: Information Canada.

Canada. Statistics Canada. 1975. 1971 Census of Canada, Industries: Employed Labour Force by Industry, Age and Sex, for Canada and Provinces, 1971. Catalogue No. 94-747, Vol. 3 - Part 4 (Bulletin 3. 4-10). Ottawa: Information Canada.

Canada. Statistics Canada. 1976. Earnings and Work Histories of the 1972 Canadian Labour Force. Catalogue No. 13-557 occasional. Consumer Income and Expenditure Division.

Canada. Statistics Canada. 1978. Advance Statistics of Education. Education Series, Catalogue No. 81-220.

Ghez, G. R., and Becker, G. S. 1975. The Allocation of Time and Goods Over Time. New York: Columbia University Press.

Gunderson, M. 1976. Work Patterns, in G.C.A. Cook, ed., Opportunity for Choice. Ottawa: Statistics Canada, pp. 93-142.

Haley, J.W. 1973. Human Capital: the Choice Between Investment and Income. American Economic Review 63: pp. 929-44.

Heckman, J. J. 1976a. A Life Cycle Model of Earnings, Learning, and Consumption. Journal of Political Economy 84: pp. S11-S44 (supplement).

Heckman, J. J. 1976b. The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models. Annals of Social and Economic Measurement 5: pp. 475-92.

Holmes, R. A. 1976. Male-female Earnings Differentials in Canada. Journal of Human Resources 11: pp. 109-17.

Irvine, I. 1978. The Distribution of Economic Welfare In Canada. Unpublished Ph.D. Thesis, University of Western Ontario, London, Canada.

Lydall, H. 1968. The Structure of Earnings. Oxford: The Clarendon Press.

Masters, S. 1974. The Effect of Educational Differences and Labor-Market Discrimination on the Relative Earnings of Black Males: Journal of Human Resources 9: pp. 342-60.

Metcalf, C. E. 1972. An Econometric Model of the Income Distribution. Chicago: Markham Publishing Co.

Mincer, J. 1974. Schooling, Experience, and Earnings. New York: Columbia University Press.

Mincer, J., and Polachek, S. 1974. Family Investment in Human Capital. Journal of Political Economy 82: pp. S76-S108 (supplement).

Oaxaca, R. 1973. Male-female Wage Differentials in Urban Labor Markets. International Economic Review 14: pp. 693-709.

Osberg, L. 1975. A Structural Approach to the Distribution of Earnings. Unpublished Ph.D. dissertation, Yale University.

Podoluk, J. R. 1968. Incomes of Canadians. Ottawa: Dominion Bureau of Statistics.

Porter, J. 1965. The Vertical Mosaic. Toronto: University of Toronto Press.

Rashid, A. 1976. An Evaluation of Income Data from the 1971 Census of Canada. Characteristics Division, Research Memorandum No. 71-EC-5.

Raynauld, A., Marion, G. and Béland, R. 1966. La répartition des revenus selon les groupes ethniques au Canada. 3 vols. Study prepared for the Royal Commission on Bilingualism and Biculturalism.

BIBLIOGRAPHY - Concluded

Robb, R. E. 1978. Earnings Differentials Between Males and Females in Ontario, 1971. Canadian Journal of Economics 11: pp. 350-359.

Rosen, S. 1977. Human Capital: Relations Between Education and Earnings, in M. D. Intriligator and D. A. Kendrick, eds., Frontiers in Quantitative Economics Vol. IIIB. Amsterdam: North Holland.

Singh, B., Nagar, A L., Choudry, N. K., and Raj.B. 1976. On the Estimation of Structural Change: a Generalization of the Random Coefficients Model. International Economic Review 17: pp. 340-61.

Sjaastad, L. A. 1962. The Costs and Returns of Human Migration. Journal of Political Economy 60: pp. 80-93.

Thurow, L.C. 1970. Analyzing the American Income Distribution. American Economic Review 60: pp. 261-9.

Tinbergen, J. 1975. Income Distribution Analysis and Policy. Amsterdam: North-Holland Press.

Wales, T. J., and Woodland, A. D. 1977. Sample Selectivity and the Estimation of Labour Supply Functions. Discussion Paper No. 77-17, Department of Economics, University of British Columbia.